

**UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
CENTRO DE CIÊNCIAS HUMANAS E SOCIAIS
DEPARTAMENTO DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO
MESTRADO EM ESTUDOS DE LINGUAGENS**

DANIELA DE SOUZA SILVA COSTA

**O LÉXICO INDÍGENA NAS CAPITAIS BRASILEIRAS: UM ESTUDO
GEOLINGUÍSTICO**

Campo Grande – MS
Dezembro-2012

DANIELA DE SOUZA SILVA COSTA

**O LÉXICO INDÍGENA NAS CAPITAIS BRASILEIRAS: UM ESTUDO
GEOLINGUÍSTICO**

Dissertação apresentada para obtenção do título de Mestre ao Programa de Pós-Graduação em Estudos de Linguagens, da Universidade Federal de Mato Grosso do Sul, sob a orientação da Prof^a Dr^a Aparecida Negri Isquardo.
Área de Concentração: Linguística e Semiótica

Campo Grande – MS
Dezembro-2012

DANIELA DE SOUZA SILVA COSTA

**O LÉXICO INDÍGENA NAS CAPITAIS BRASILEIRAS: UM ESTUDO
GEOLINGUÍSTICO**

APROVADA POR:

APARECIDA NEGRI ISQUERDO, DOUTORA (UFMS)

AURI CLAUDIONEI MATOS FRÜBEL, DOUTOR (UFMS)

SUZANA ALICE MARCELINO CARDOSO, DOUTORA (UFBA)

Campo Grande, MS, 06 de dezembro de 2012.

DEDICATÓRIA

*Aos amores da minha vida
Érico e Gabriela*

AGRADECIMENTOS

A Deus, todo-poderoso, que me concedeu a graça da vida e todas as bênçãos dia a dia, e especialmente por me munir de tudo o que precisava para concluir esta jornada.

Ao meu marido, Érico, companheiro de uma vida, sempre solícito, disponível, amigo, e sempre com uma palavra de estímulo nas horas mais difíceis. Partilhar esta conquista com você só me dá a certeza de que Deus escolheu o melhor para mim e que sou muito feliz ao seu lado.

Ao meu pedaço de céu, Gabriela, a enviada por Deus e a maior graça a nós concedida, este trabalho e toda a minha vida passaram a valer mais a partir da chegada dela.

A minha família, que me acolheu em todos os momentos, nos de alegria, de tristeza, de insegurança, de medos e hesitações, agradeço a Deus por ter me concedido pessoas tão maravilhosas que vieram por laços de sangue e de amor.

Aos amigos que estiveram presentes em todos os momentos ou em alguns deles, sempre me encorajando, aos amigos antigos, aos novos, aos virtuais, a todos, meu muito obrigada.

A minha querida orientadora e amiga Profa. Dra. Aparecida Negri Isquerdo, por tudo de bom que me ofereceu nesta caminhada: pelas orientações, pelos conselhos, mas principalmente pela amizade, pelo carinho e pela dedicação que sempre me ofereceu. É uma honra poder compartilhar momentos com você, mestra, que me ensinou a amar a pesquisa desde os tempos de graduação, mas

que essencialmente me ajudou a me tornar uma pessoa melhor, por seu exemplo de conduta diante da vida.

A todos os professores do Mestrado, que ofereceram seu tempo para me fornecer subsídios para concluir a Pós-Graduação, tanto no que diz respeito às aulas quanto às conversas, conselhos e amizade oferecida. Também tenho muito a agradecer aos servidores que atuam e atuaram no Mestrado em Estudos de Linguagens, especialmente as secretárias Daniela e Ana Carla, sempre muito prestativas, atenciosas e generosas.

Aos meus amigos do Mestrado, a quem dedico um parágrafo a parte. Vocês me ensinaram que a Pós-Graduação também é lugar de troca, de generosidade, de amizade, e por que não dizer também de diversão: as conversas nos intervalos, as reuniões extras, a troca de materiais e o compartilhamento de e-mails, cópias de textos, auxílio virtual via Hotmail, Facebook, Dropbox e o que mais a tecnologia pode nos oferecer para diminuir distâncias. Levarei vocês comigo no coração, sempre.

Às amigas negretes, orientandas da Profa. Aparecida. Amo fazer parte dessa família, sempre unida, amiga, prestativa e pronta para ajudar, especialmente a Luciene, negrete que se tornou afilhada e uma amiga muito especial, pela ajuda com as cartas linguísticas, com os conselhos e a sobriedade que tem quem tem Deus no coração.

A todos os que cruzaram minha vida durante este caminhar e mesmo quando ele era apenas um sonho, da graduação até este momento, a todos os que não consegui nomear aqui ou que talvez possam não se encaixar em algum dos agradecimentos anteriores: sou grata pela existência de cada pessoa que estive e está comigo, pois a convivência com o outro nos faz superar as

adversidades e crescer como profissionais, cristãos e peças dessa máquina chamada mundo.

RESUMO

A língua reflete influências do ambiente social, como os determinantes sócio-históricos, culturais e geográficos relacionados à história social dos falantes. E, especialmente no léxico, essas influências são percebidas, por se refletirem nas escolhas dos usuários. No Brasil, por exemplo, desde os tempos coloniais, os intercâmbios culturais e linguísticos que se processaram entre europeus, indígenas e africanos deixaram marcas e contribuíram sobremaneira para a construção da identidade nacional. Assim, no âmbito do léxico, os empréstimos lexicais advindos das línguas de povos indígenas e africanos, além das línguas de imigrantes, muito enriqueceram o português do Brasil com um vocabulário voltado para as diferentes realidades regionais do país. Este trabalho utilizou como arcabouço teórico princípios da Dialectologia, da Geolinguística, da Lexicologia e da Semântica e, como *corpus*, dados geolinguísticos coletados pela equipe de pesquisadores do Projeto Atlas Linguístico do Brasil (ALiB) nas 25 capitais brasileiras que integram a rede de pontos do Projeto ALiB. O estudo teve como propósito investigar a influência indígena no repertório lexical dos falantes dessas localidades, expressa por meio do uso de indigenismos – unidades léxicas de base indígena – relacionados a áreas semânticas como a *fauna*, as *atividades agropastoris* e a *alimentação e cozinha*. Para isso, além de se analisar os dados pesquisados sob a luz dos ramos da Linguística mencionados, buscou-se, na história social das localidades investigadas, possíveis motivações para a manutenção e/ou desaparecimento do substrato indígena no vocabulário dos habitantes de grandes centros urbanos, demonstrando também, por meio da cartografia dos dados analisados, possíveis isoglossas léxicas que demarcam áreas dialetais em território nacional. Nas 62 perguntas das três áreas semânticas pesquisadas, dos 800 itens lexicais catalogados, 59 são de base indígena (7,3%), em sua maioria no campo da *fauna*: 30 ocorrências. Com isso, identificaram-se também certas isoglossas léxicas que evidenciam traços regionais no vocabulário, como os usos de *aipim*, *muriçoca* e *curau*, por exemplo, e também que as cidades comportam-se de maneiras distintas, no que diz respeito às escolhas lexicais de seus falantes. O estudo ratificou ainda que a importância das pesquisas geolinguísticas como subsídios para dicionários gerais da língua, no que tange ao registro de marcas dialetais, à medida que revelou, ora a disseminação de unidades léxicas dicionarizadas com marca de uso de certas regiões, como *muriçoca*, ora a confirmação de marcas de uso já registradas em obras lexicográficas, como o uso de *carapanã*, e mesmo a inovação lexical promovida pelos usuários ao atribuir novas acepções a itens lexicais já existentes, como o ocorrido com *soca* e *carioquinha*. Em síntese, os dados analisados reiteraram o papel da Geolinguística para o conhecimento da língua em uso, num dado espaço e tempo, aqui representados pela contemporaneidade no uso da língua portuguesa no Brasil.

Palavras-chave: Indigenismos. Norma lexical. Capitais brasileiras.

ABSTRACT

The language reflects influences of the social environment, such as socio-historical, cultural and geographic determiners related to the social history of the speakers. And especially in the lexicon these influences are realized, being reflected on the choices of users. In Brazil, for example, since colonial times, linguistic and cultural exchanges that there were between Europeans, Indians and Africans left their marks and contributed greatly to the construction of Brazilian identity. Thus, within the lexicon, lexical loans arising from indigenous and African languages, besides languages of the immigrants, greatly enriched the Portuguese of Brazil with a vocabulary geared to the different regional realities of the country. This research used as the theoretical principles of Dialectology, Geolinguistics, Lexicology and Semantics, and as corpus, geolinguistic data collected by the research team of the Linguistic Atlas Project in Brazil (ALiB) in the 25 Brazilian capitals that integrate its point net. The study has as purpose to investigate the Indian influence on lexical repertoire speakers of Brazilian capitals, expressed through the use of indigenisms - lexical units of indigenous base – related to semantic areas dealing with national elements, such as wildlife, agropastoral activities and food and cooking. For this, besides analyzing the data surveyed in the light of science has already mentioned, we sought in the social history of the localities, possible motivations for the maintenance and / or disappearance of this substrate in the vocabulary of inhabitants of big urban centres, also demonstrating, through mapping data analyzed for possible lexical isoglosses that mark dialectal areas in the country. In the 62 questions of the three surveyed semantic areas, from 800 lexical items cataloged, 59 are of indigenous base (7, 3%), mostly in the wildlife field: 30 occurrences. Thus, we identified certain lexical isoglosses scoring uses regional vocabulary, as the usages of *aipim*, *muriçoca* and *curau*, and also that cities behave differently with regard to the lexical choices of their speakers. Moreover, the study confirmed the importance of geolinguistic research to subsidize language general dictionaries, in dialectal mark registers, as soon as revealed a spread of lexical units registered as typical of certain regions such as *muriçoca*, or the confirmation of usage marks registered, such as using *carapanã*, and even the lexical innovation promoted by users to attribute new meanings to existing lexical items, such as occurred with *soca* and *carioquina*. In short, data reiterated the rule of Geolinguistics to the knowledge of the language in use at a given time and space, here represented by the contemporary use of Portuguese in Brazil.

Keywords: Indigenisms. Lexical Standard. Brazilian capitals.

LISTA DE MAPAS

| Título | Página |
|--|--------|
| Mapa 1 - Mapa etnolinguístico da realidade indígena no Brasil segundo Mattoso Câmara (1977) | 29 |
| Mapa 2 - Etnias indígenas do Brasil segundo Bethell (2004) | 34 |
| Mapa 3 - Rotas das Bandeiras e seus objetivos – séculos XVI a XIX | 44 |
| Carta Linguística 1 – Indigenismos – Produtividade geral | 109 |
| Carta Linguística 2 – <i>Mandioca, macaxeira e aipim</i> – Distribuição diatópica | 128 |
| Carta Linguística 3 – Indigenismos mais produtivos – <i>Atividades agropastoris</i> – Distribuição diatópica | 134 |
| Carta Linguística 4 – Pernilongo – Base linguística | 159 |
| Carta Linguística 5 – <i>Gambá</i> – Distribuição diatópica | 168 |
| Carta Linguística 6 – <i>Muriçoca</i> – Produtividade | 171 |
| Carta Linguística 7 – Isoglossas léxicas – fauna | 173 |
| Carta Linguística 8 – Indigenismos – <i>Alimentação e cozinha</i> – Distribuição diatópica | 184 |
| Carta Linguística 9 – <i>Mingau</i> – Distribuição diatópica | 186 |
| Carta Linguística 10 – <i>Curau/curau de milho</i> – Produtividade | 188 |
| Carta Linguística 11 – <i>Pão carioca/carioquinha</i> – produtividade | 190 |

LISTA DE QUADROS

| Título | Página |
|--|--------|
| Quadro 1 – Questionário Semântico-lexical/ALiB, área semântica <i>atividades agropastoris</i> | 87 |
| Quadro 2 – Questionário Semântico-lexical/ALiB, área semântica <i>fauna</i> | 89 |
| Quadro 3 – Questionário Semântico-lexical/ALiB, área semântica <i>alimentação e cozinha</i> | 91 |
| Quadro 4 - Indigenismos documentados na área semântica <i>atividades agropastoris</i> nas capitais brasileiras | 110 |
| Quadro 5 – Respostas não válidas em relação ao referente requerido pela pergunta na área semântica <i>atividades agropastoris</i> | 113 |
| Quadro 6 – Distribuição dos indigenismos na área semântica das <i>atividades agropastoris</i> documentados nas capitais brasileiras de acordo com a região do Brasil | 123 |
| Quadro 7 - Produtividade dos indigenismos registrados na área semântica das <i>atividades agropastoris</i> nas capitais brasileiras | 124 |
| Quadro 8 - Percentuais de ocorrência de mandioca, macaxeira e aipim nas capitais brasileiras | 129 |
| Quadro 9 – Marcas dialetais dicionarizadas e documentadas pelo Projeto ALiB - itens lexicais da área semântica das <i>atividades agropastoris</i> | 136 |
| Quadro 10 - Indigenismos documentados na área semântica da <i>fauna</i> , de acordo com a pergunta do Questionário Semântico-lexical/ALiB | 138 |
| Quadro 11 - Traços semânticos das unidades lexicais <i>gambá</i> e <i>jaratataca</i> | 154 |
| Quadro 12 – Distribuição dos tupinismos na área semântica da <i>fauna</i> , documentados nas capitais brasileiras de acordo com a região do Brasil | 163 |
| Quadro 13 – Marcas dialetais dicionarizadas e documentadas pelo Projeto ALiB - itens lexicais da área semântica da <i>fauna</i> | 175 |
| Quadro 14 - Indigenismos documentados na área semântica <i>alimentação e cozinha</i> nas capitais brasileiras | 176 |
| Quadro 15 – Marcas dialetais dicionarizadas e documentadas pelo Projeto ALiB - itens lexicais da área semântica da <i>alimentação e cozinha</i> | 192 |

LISTA DE FIGURAS

| Título | Página |
|--|---------------|
| Figura 1 – Tela inicial do SGP | 93 |
| Figura 2 – Cadastro de perguntas no SGP | 94 |
| Figura 3 – Cadastramento inicial no SGP | 95 |
| Figura 4 – Cadastro de resposta no sistema SGP | 95 |
| Figura 5 – Tela de cadastros do SGP | 96 |
| Figura 6 – Tela de classificação de respostas do SGP | 97 |
| Figura 7 – Listagem de respostas por língua de origem | 98 |
| Figura 8 – Listagem diatópica de indigenismos segundo a capital brasileira | 99 |

LISTA DE GRÁFICOS

| | |
|--|-----|
| Gráfico 1 – Distribuição do total dos itens lexicais documentados de acordo com a base linguística | 103 |
| Gráfico 2 – Indigenismos | 106 |
| Gráfico 3 – Distribuição das unidades lexicais documentadas na área semântica das <i>atividades agropastoris</i> de acordo com a base linguística | 112 |
| Gráfico 4 – Distribuição diageracional de <i>jacá</i> nas capitais brasileiras | 132 |
| Gráfico 5 – Distribuição dos itens lexicais documentados na área semântica da <i>fauna</i> de acordo com a base linguística nas capitais brasileiras | 138 |
| Gráfico 6 – Produtividade de <i>mucura</i> nas capitais brasileiras | 169 |
| Gráfico 7 – Distribuição das unidades lexicais relacionadas à área semântica da <i>alimentação e cozinha</i> de acordo com a base linguística | 177 |

LISTA DE ABREVIATURAS E SIGLAS

AL – Alagoas

ALiB – Atlas Linguístico do Brasil

AM – Amazonas

CE – Ceará

Cm – centímetro

D – dom

ES – Espírito Santo

Etc. – *et cetera*

IBGE – Instituto Brasileiro de Geografia e Estatística

INF – informante

INQ – inquiridor

LaLI – Laboratório de Línguas Indígenas

LG – Língua Geral

Mm – milímetro

MS – Mato Grosso do Sul

PB – Português do Brasil

PE – Pernambuco

PI – Piauí

QSL – Questionário Semântico-Lexical

RN – Rio Grande do Norte

RO – Rondônia

RR – Roraima

SC – Santa Catarina

SGP – Sistema Geral de Pesquisa

Sic – assim

Sin. – sinônimo

SP – São Paulo

UnB – Universidade de Brasília

SUMÁRIO

| | |
|--|----|
| Introdução | 16 |
| PARTE I | |
| CAPÍTULO I – 500 ANOS DE BRASIL? CONTEXTUALIZAÇÃO SÓCIO-HISTÓRICA DO PAÍS DE BASES INDÍGENAS | 22 |
| 1.1. Os 500 anos de domínio português | 22 |
| 1.1.1. Período colonial: aspectos da sociedade, da língua e da economia | 25 |
| 1.1.1.1. O ciclo da cana-de-açúcar e a estruturação da economia colonial | 37 |
| 1.1.1.2. O ciclo do ouro e a alteração de polos no Brasil Colônia | 41 |
| 1.1.1.3. A crise no sistema colonial | 47 |
| 1.1.2. Período pós-colonial: o estabelecimento da soberania nacional | 48 |
| 1.1.2.1. O ciclo do café e a extração ervateira: um novo panorama étnico no Brasil | 50 |
| CAPÍTULO II – PRESSUPOSTOS TEÓRICOS: O LÉXICO EM FOCO | 53 |
| 2.1. O conceito de norma lexical e sua aplicabilidade para o estudo de indigenismos no léxico do português do Brasil | 55 |
| 2.2. Apresentação do aporte teórico | 62 |
| 2.3. O português do Brasil: achegas a um estudo etnolinguístico | 70 |
| 2.4. O legado indígena no léxico do Brasil: da Língua Geral aos dias atuais | 77 |

PARTE II

| | |
|---|-----|
| CAPÍTULO III – METODOLOGIA DA PESQUISA | 81 |
| 3.1. Projeto Atlas Linguístico do Brasil: revelando falares, descortinando a língua em uso | 83 |
| 3.2. O SGP – Sistema Geral de Pesquisa: a tecnologia da informação auxiliando a pesquisa linguística | 92 |
| 3.3. Tratamento dos dados | 99 |
| | |
| CAPÍTULO IV – APRESENTAÇÃO E ANÁLISE DOS DADOS | 102 |
| 4.1 Análise de dados da área semântica <i>atividades agropastoris</i> | 110 |
| 4.2. Análise de dados da área semântica <i>fauna</i> | 137 |
| 4.3. Análise de dados da área semântica <i>alimentação e cozinha</i> | 176 |
| CONSIDERAÇÕES FINAIS | 194 |
| REFERÊNCIAS | 202 |

INTRODUÇÃO

O léxico é o acervo vocabular de uma língua, utilizado pelo homem para nomear a realidade ao seu redor, transmitir ideias e mesmo para organizar pensamentos. Esse repertório lexical reflete aspectos culturais da sociedade que o utiliza e seu estudo propicia não só um melhor conhecimento não só da língua em uso por um dado grupo social, como também fornece pistas sobre o modo de viver e de pensar desse mesmo grupo, além de propiciar subsídios para a identificação de influências interétnicas incorporadas pelo acervo lexical de uma língua.

Nesse sentido, contatos entre diversos povos refletem-se na língua em uso por um determinado grupo social. No Brasil, por exemplo, a variante brasileira do português desenhou-se com particularidades frente à língua transplantada pelo colonizador, uma vez que o europeu aqui se encontrou com outras etnias que contribuíram sobremaneira para a constituição do repertório lexical dos habitantes da nova nação em construção.

Isso porque o léxico reflete “em alto grau a complexidade da cultura [...] destina-se em qualquer época a funcionar como um conjunto de símbolos, referentes ao quadro cultural do grupo” (SAPIR, 1969, p. 51). No caso do Brasil, esse quadro social é delineado a partir dos elementos indígena, branco e negro, miscigenados e interagentes na formação da identidade nacional.

Os contatos interétnicos iniciais, no Brasil, ocorreram no primeiro decanato do século XVI, com a chegada dos portugueses. Encantados com uma realidade totalmente diferente da europeia, já que o novo mundo abrigava matas

fechadas, grandes árvores e animais nunca vistos, os europeus encontraram na costa brasileira a maior etnia indígena habitante dessas terras: os índios tupi. Sabe-se que havia, nas terras de Vera Cruz, à época do descobrimento, diversas etnias indígenas que falavam mais de mil línguas (RODRIGUES, 2002); porém, o encontro com os tupi deveu-se especialmente ao fato de essa nação habitar o litoral brasileiro, portal de entrada dos europeus em terras nacionais.

Inicialmente, os portugueses que vinham para a Colônia valiam-se da mão de obra indígena para a extração do pau-brasil, haja vista que os tupis detinham o conhecimento dessa atividade. Nesse contexto, Fernandes (1985, p.80) esclarece que

[...] O estudo da evolução da situação de contacto põe em evidência as condições dentro das quais o sistema organizatório tribal podia reagir construtivamente à presença dos brancos. Enquanto estes eram em pequeno número e podiam ser incorporados à vida social aborígine ou se acomodavam às exigências dela, nada afetou a unidade e a autonomia do sistema social tribal. Essa situação manteve-se onde os brancos se limitavam à exploração de produtos que podiam ser permutados com os índios, especialmente o pau-brasil. O intercâmbio econômico, nessas condições, não exigia a permanência de grande número de estrangeiros nos grupos locais, dependendo tanto sua alimentação quanto sua segurança do que decidiam fazer os “aliados” indígenas.

Estreitando-se os laços dos brancos com os autóctones, também os casamentos entre índias e portugueses passaram a ser benquistos na sociedade indígena, ocasionando a miscigenação tão característica da nação brasileira. Isso originou famílias híbridas, falantes, no seio familiar, de uma língua de base indígena, ficando o uso do português restrito aos documentos oficiais e ao ensino escolar.

Assim, essa língua híbrida se constituiu na forma comum de comunicação na Colônia Portuguesa, a Língua Geral¹ (LG), uma vez que a diversidade étnica

¹ Três eram as línguas gerais do período colonial: a Língua Geral Paulista (LGP), a Língua Geral Amazônica (LGA) e o Guaraní Criollo (GNC) (RODRIGUES, 2012b). Utilizamos-nos da sigla LG para designar a Língua Geral Paulista, derivada do tupinambá ou tupi antigo e falada pelos bandeirantes

favoreceu essa miscigenação linguística, dando origem a essa língua de base indígena com inserções de unidades léxicas portuguesas.

Dessa forma, a Colônia passou a se comunicar por meio da Língua Geral, domínio esse que se espalhou por todo o território também por obra das bandeiras paulistas, que reunia também índios e mestiços falantes da LG.

Esse predomínio da LG durou aproximadamente três séculos². Mesmo assim, depois da proibição de seu uso, o português já havia sido influenciado pela língua de base indígena, o que se refletiu em diversos empréstimos linguísticos de base autóctone presentes na norma lexical da variante brasileira do português.

Além disso, apesar do predomínio do tupi no cenário indígena nacional, dadas a população dos índios dessa etnia e também a história de contatos entre as diversas nações indígenas, outros índios também fizeram e fazem parte da constituição da identidade brasileira, mesmo que em menor proporção.

Nesse sentido, itens lexicais de base indígena ainda integram o léxico do português do Brasil e, por isso, este trabalho visa a identificar a vitalidade dessas unidades na língua falada pelos habitantes das capitais brasileiras que integram a rede de pontos do Projeto Atlas Linguístico do Brasil (Projeto ALiB).

Assim, tendo em vista a relevância do elemento indígena na constituição da identidade brasileira e a permanência de sua influência no português do Brasil, este trabalho investigou a questão da manutenção de indigenismos no léxico do português do Brasil, atualizado na fala dos habitantes de capitais brasileiras³, além

paulistas e em todo o centro-sul do Brasil à época, que interagiu com o português europeu e forneceu parte de seu vocabulário para a formação da variante brasileira do Português. A Língua Geral amazônica, por seu turno, era utilizada na região do atual Estado do Amazonas e o Guarani Criollo, por sua vez, concentrou-se a oeste do atual Estado do Paraná, no Brasil.

² O predomínio da Língua Geral durou dos séculos XVI até meados do século XVIII, isto é, desde o início da colonização até 1759, quando Marquês de Pombal proibiu a sua utilização.

³ Excluíram-se as capitais de Tocantins (Palmas) e do Distrito Federal (Brasília), haja vista que, por serem de fundação recente à época da definição da rede de pontos do Projeto ALiB (final da década

de analisar em que medida motivações extralinguísticas, como fatores sócio-histórico-culturais, podem auxiliar na manutenção ou no desaparecimento da presença indígena na identidade brasileira, aqui representada pelo léxico do português do Brasil, modalidade oral.

Para se alcançar esse objetivo geral, traçaram-se também os seguintes objetivos específicos:

- Analisar o recorte de unidades lexicais de base indígena que nomeiam referentes relacionados às áreas semânticas da *fauna*, da *alimentação e cozinha* e das *atividades agropastoris*, documentadas pelo Projeto ALiB na fala de habitantes de capitais brasileiras, sob as perspectivas diatópica e léxico-semântica.
- Verificar em que proporção a presença direta e/ou indireta de etnias indígenas nas localidades pesquisadas reflete-se na manutenção ou não de indigenismos na língua oral dos habitantes das capitais brasileiras.
- Demonstrar a distribuição diatópica dos itens lexicais documentados, buscando na história social, na formação étnica da população e nas características histórico-ambientais das localidades pesquisadas possíveis motivações para as designações registradas.
- Verificar a existência de possíveis isoglossas léxicas em território nacional, por meio da investigação de dados lexicais coletados nas capitais brasileiras.
- Fornecer contribuições para as pesquisas do Atlas Linguístico do Brasil, bem como para outros estudos acadêmicos da área, dentre outros, a Lexicografia (fornecimento de fontes para os dicionários em termos de marcas de uso).

Para isso, foram analisados dados já coletados por pesquisadores brasileiros das diferentes regiões do País para a composição de um atlas linguístico nacional, o Atlas Linguístico do Brasil – Projeto ALiB, referentes a áreas semânticas selecionadas que tendem a contemplar de forma mais pontual elementos nacionais, como as *atividades agropastoris*, a *fauna* e a *alimentação e cozinha*, na expectativa de que revelassem grande contingente de unidades léxicas de origem indígena.

E, com vistas a uma melhor visualização do contexto da pesquisa, esta Dissertação foi estruturada em duas partes. Na primeira, além dos elementos pré-textuais, apresenta-se o Capítulo I, que versa sobre aspectos da história do Brasil, especialmente no que diz respeito aos contatos interétnicos ocorridos em terras brasileiras, suas motivações e consequências linguísticas, econômicas, sociais, históricas e políticas.

Depois, no Capítulo II, apresenta-se o aporte teórico que embasou a pesquisa. Nessa parte, são discutidos aspectos relativos à norma lexical e aos trabalhos relacionados ao léxico propriamente ditos, além de serem focalizados aspectos das ciências que deram base a este trabalho, além de um estudo do português do Brasil sob um viés etnolinguístico.

Realizada, pois, a contextualização em sua face histórica e teórica, a segunda parte do trabalho traz, de fato, os resultados do estudo em questão, por meio de seus Capítulos III e IV.

No Capítulo III, é apresentado o percurso metodológico tomado, além de ser demonstrado o programa de computador que auxiliou o armazenamento, a consulta e a análise dos dados documentados. Também esse Capítulo traz informações sobre o Projeto Atlas Linguístico do Brasil, a que se vincula este trabalho.

Por fim, o Capítulo IV apresenta a análise dos dados lexicais vinculados às três áreas semânticas selecionadas para o trabalho. Aqui, são apresentados gráficos e cartas linguísticas que ilustram as análises, bem como o cotejo diatópico e léxico-semântico das unidades lexicais documentadas.

Concluindo este trabalho, encontram-se as considerações finais acerca da natureza dos dados analisados, bem como as referências consultadas.

CAPÍTULO I

500 ANOS DE BRASIL? CONTEXTUALIZAÇÃO SÓCIO-HISTÓRICA DO PAÍS DE BASES INDÍGENAS

Quando os portugueses chegaram ao Brasil, há mais de quinhentos anos, como já assinalado, aqui se encontraram com diversos povos, falantes de distintas línguas indígenas (RODRIGUES, 2002). Essas sociedades tinham sua organização e suas tradições já sedimentadas de acordo com suas crenças e vivências, o que fez com que o europeu buscasse a cooperação com esses povos.

E essa relação, entre índios e europeus, foi marcante na formação da sociedade colonial brasileira, que depois recebeu os negros africanos e imigrantes de outros países. A base da formação étnica do povo brasileiro ancorou-se em três etnias: a branca (portuguesa), a indígena e a africana, cuja miscigenação gerou a identidade linguística e social do Brasil, que se reflete até os dias atuais.

Aspectos desse momento histórico e social foram abordados neste Capítulo, que trata da relação entre índios, brancos e negros, bem como da formação político-social e econômica do Brasil, do Período Colonial até a era republicana.

1.1 OS QUINHENTOS ANOS DE DOMÍNIO PORTUGUÊS

[...] A terra do Brasil, que está na América, uma das quatro partes do mundo, não se descobriu de propósito, e de principal intento; mas

acaso indo Pedro Álvares Cabral, por mandado de El-rei d. Manuel, no ano de 1500 para as Índias, por capitão-mor de 12 naus, afastando-se da costa de Guiné, que já era descoberta ao Oriente, achou estoura ao Ocidente, da qual não havia notícia alguma, foi costeando alguns dias com tormenta até chegar a um porto seguro, do qual a terra vizinha ficou com o mesmo nome. Ali desembarcou o dito capitão com seus soldados armados, para pelejarem; porque mandou primeiro um batel com alguns a descobrir campo, e deram novas de muitos gentios, que viram; porém não foram necessárias armas, porque só de verem homens vestidos, e calçados, brancos, e com barba / do que tudo eles carecem / os tiveram por divinos, e mais que homens, e assim chamando-lhe Caraíbas, que quer dizer na sua língua coisa divina, se chegaram pacificamente aos nossos (SALVADOR, 1627, p. 03).

O relato de Frei Vicente Salvador, mais de um século após a esquadra de Cabral aportar em terras brasileiras, simboliza a visão do colonizador acerca da nova terra recém-descoberta, bem como a primeira impressão acerca dos povos nativos. Além disso, representa a corrida expansionista por que passou, especialmente, a Península Ibérica, no século XV, com vistas a descobrir novas praças fornecedoras de matérias-primas, assim como essa cruzada expansionista possuía também o intuito da defesa da fé cristã propagada pelo catolicismo a todos os povos.

E a Terra de Vera Cruz assim tornou-se território português além-mar. Esse nome deveu-se ao fato de Pedro Álvares Cabral ter aqui levantado uma cruz, exatamente simbolizando o domínio da religião católica, em 03 de maio, dia em que, no calendário português, comemorava-se a invenção da Santa Cruz.

A rota para as Índias a partir do Ocidente foi proposta, na verdade, por Cristóvão Colombo, que chegou às Grandes Antilhas em 1492, descobrindo o Novo Mundo a serviço dos espanhóis⁴. Dessa forma, os portugueses seguiram também pelo Ocidente para alcançar as Índias, produtoras de especiarias que tanto agradavam o mercado consumidor europeu, chegando assim ao território que viria a

⁴ Há vertentes de estudos que atribuem a descoberta do Novo Mundo a outros desbravadores. Todavia, pelo fato de esse assunto não ser o tema principal deste trabalho, optou-se por adotar a versão mais difundida do tema, sem maiores discussões.

ser identificado como Brasil. Na esquadra de Cabral, assim como em navios que depois aportaram na então Colônia Portuguesa, os tripulantes eram, em sua quase totalidade, homens que buscavam novos rumos para suas vidas, estavam à procura de riquezas, fortuna e novas perspectivas de vida.

Segundo Holanda (1988, p.14), a característica principal desse homem europeu que se tornou colonizador de terras brasileiras era o espírito aventureiro. Essa visão justifica dois pontos centrais para a Colonização Portuguesa no Brasil: a adaptação do português às novas terras, às riquezas e aos produtos passíveis de exploração – ao invés de implantar seus modos de produção e mesmo seu modo de viver, ajustou-se às características do solo brasileiro e a seus produtos – e a inserção nas sociedades indígenas sem promover mudanças significativas em sua organização, motivo pelo qual o branco era bem recebido pelos índios, sem a necessidade de peleja, como citado por Salvador (1627) no relato que inicia este Capítulo.

Nesse particular, Fernandes (1985, p. 74) acrescenta que

[...] Enquanto estes [os brancos] eram em pequeno número e podiam se incorporados à vida social aborígine ou se acomodavam às exigências dela, nada afetou a unidade e a autonomia do sistema social tribal. Essa situação manteve-se onde os brancos se limitavam à exploração de produtos que podiam ser permutados com os índios, especialmente o pau-brasil. O intercâmbio econômico, nessas condições, não exigia a permanência de grande número de estrangeiros nos grupos locais, dependendo tanto sua alimentação quanto sua segurança do que decidiam fazer os “aliados” indígenas.

Porém, apesar de estarem em menor número, os colonizadores acabavam por subjugar os indígenas, uma vez que lhes traziam uma realidade completamente desconhecida e encantadora a seus olhos. Eles usavam vestes, por exemplo, e estas eram coloridas e ricas em detalhes. Também tinham a pele alva, os olhos claros e usavam barba, características nunca antes vistas pelos

autóctones, isso tudo associado aos objetos que traziam consigo. Eram espelhos, armas e artefatos de metal reluzente, por exemplo, que encantavam os povos “pardos, nus, sem coisa alguma que lhes cobrisse as vergonhas. [e que] Traziam arcos nas mãos, e suas setas” (CAMINHA, 1963 [1500], p. 03).

Assim, aos poucos, o homem branco foi alçando *status* superior na sociedade que estava se reestruturando por sua presença e, conhecendo melhor as riquezas desta terra e almejando lucros advindos de sua exploração, relatava à Coroa Portuguesa todas as descobertas e perspectivas de vantagens para a economia real, o que aumentou o interesse de Portugal pelo Brasil e o fez assumir a posse da nova terra a partir de 1532, tornando-se, de fato, Colônia Portuguesa.

1.1.1 Período colonial: aspectos da sociedade, da língua e da economia.

O Período Colonial brasileiro pode ser dividido em três momentos distintos, mais de acordo com os acontecimentos registrados e sua importância que com o intervalo de tempo decorrido. Fausto (2008, p.41) assim delimita esses momentos: um “primeiro [que] vai da chegada de Cabral à instalação do governo geral, em 1549; o segundo [...] entre a instalação do governo geral e as últimas décadas do século XVIII; o terceiro [que] vai dessa época à Independência, em 1822”. Segundo o historiador, o primeiro momento poderia ser resumido como a descoberta e a tomada de posse do novo território. Já o segundo período diria respeito à consolidação da colônia e o terceiro concluiria a era colonial, considerando-se aspectos internos e externos da economia até a instalação da Monarquia.

Segundo a periodização proposta por Fausto (2008), o momento iniciado pelo Descobrimento do Brasil foi o primeiro passo para o reconhecimento mundial das terras brasileiras, assim como sua inserção na economia portuguesa.

À época, Portugal já se destacava como potência expansionista, pelos conhecimentos cosmográficos e matemáticos promovidos pelo infante D. Henrique a partir do século XV, que teria fundado, inclusive, uma escola náutica, a Escola de Sagres⁵.

Assim, disseminada a notícia de um possível caminho a partir do Oceano Atlântico para as Índias, descoberto por Cristóvão Colombo, os portugueses lançaram-se à empresa descobridora de novas terras, produtoras de suprimentos para atender seu mercado consumidor, cada vez mais exigente. Dessa forma, em 21 de abril de 1500, as naus capitaneadas pelo navegador Pedro Álvares Cabral chegaram às terras que acreditaram ser inicialmente uma ilha, tendo avistado

[...] primeiramente, “um grande monte, muito alto e redondo”, ao qual pôs o capitão-mor o nome de Monte Pascoal, por estarem então na Páscoa; “e outras terras mais baixas, ao lado sul dele”; e ‘terra chã [sic] com grandes arvoredos’, a que denominaram Vera Cruz (VIANNA, 1965a, p.42).

Uma cruz foi posta no Monte Pascoal para simbolizar a tomada de posse do território pelos cristãos portugueses, atitude seguida pela celebração de uma missa para os navegantes e também para os índios que se aproximavam com curiosidade. E esse território, que depois receberia o nome de Brasil, tornou-se de domínio português, bem como tudo o que nele havia: desde as riquezas naturais até seu elemento humano.

⁵ A real existência da Escola de Sagres tem sido tema de profundas discussões no campo da História. Porém, independente de sua existência ou não, notável foi o destaque da nação portuguesa nas descobertas de novas terras promovidas pela habilidade e pelos conhecimentos de seus representantes, como Vasco da Gama e Pedro Álvares Cabral, por exemplo.

Os índios de diversas etnias foram, em sua maioria, receptivos com os portugueses, tendo em vista a prática institucionalizada do *cunhadismo*. Segundo Ribeiro (1995), a formação do povo brasileiro deu-se especialmente por esse hábito indígena, que incorporava estrangeiros à sua comunidade por meio do casamento com uma índia. Dessa maneira, “aceitando a moça, o estrangeiro passava a ter nela sua *temericó*⁶” e assim estabeleceria “mil laços que o aparentavam com todos os membros do grupo” (RIBEIRO, 1995, p. 81), seja com os parentes de gerações superiores, de sua própria geração e também com a geração inferior.

Tal prática foi extremamente proveitosa para o colonizador, haja vista que poderia, assim, contar com uma “família” para de fato dominar e conhecer o universo a seu redor. E foi justamente sobre essa estruturação social que se fundou o primeiro ciclo econômico da Colônia, a extração do pau-brasil. Os índios, conhecedores de seu *habitat* e “parentes” do colonizador auxiliavam-no na busca pela madeira nobre para tingimento de tecidos e fabricação de móveis, uma vez que a árvore, apesar de abundante na mata atlântica, encontrava-se dispersa na mata, e só quem a conhecia poderia identificá-la e extraí-la. Então, durante muito tempo, estima-se que entre 1500 e 1535, o extrativismo do pau-brasil foi a principal fonte de renda e de interesse da Coroa Portuguesa no Brasil Colônia.

Além disso, os casamentos entre brancos e índias, além de benquistos pela sociedade indígena, passaram também a ser incentivados pela Coroa Portuguesa, uma vez que essa prática auxiliaria a efetiva colonização do território. Então o índio foi tomado não apenas como mão de obra para atividades extrativistas ou como escravo, mas também como participante na colonização. “Os colonos viam nele um *trabalhador* aproveitável; a metrópole, um *povoador* para a área imensa que

⁶ Termo que equivale à esposa, que, além de se unir a seu esposo, também unia toda a sua família, ascendentes e descendentes, a esse marido. A figura da *temericó* era a base do *cunhadismo*.

tinha que ocupar, muito além de sua capacidade demográfica” (PRADO JR, 1986, p. 91).

Segundo Ribeiro (1995, p. 320), a população do Brasil Colônia seria dividida em três classes, todas elas com base no elemento indígena: o índio tribal, encerrado ainda na mata fechada e alheio à colonização; a população urbanizada, heterogênea que congregava os imigrantes aqui chegados e ainda índios e depois negros, e os índios genéricos, assim chamados por serem mamelucos descendentes das uniões conjugais e extraconjugais entre brancos e índias.

Ainda sobre esse caldeirão de etnias, convém salientarem-se as características linguísticas desse embrião da população brasileira, uma vez que o colonizador trouxe o português, que aqui se pôs em contato com uma extensa variedade de línguas indígenas ao longo do tempo. Dentre essas línguas, o fato de haver um sem número de referências ao domínio da língua indígena tupi durante a colonização brasileira deve-se a certos fatores, incluindo-se, claro, a hegemonia dessa etnia sobre as demais, além de revelar a visão do colonizador frente à sociedade em formação. Segundo Borges (2001, p. 215),

[...] No que concerne ao imaginário e à memória linguística brasileiros, há uma recorrência da imagem de uma língua comum a todos os índios. A implicação disso é que subjaz à concepção de uma categoria genérica – o índio – que indiferencia todos os grupos indígenas. Esta categoria uniformizadora pressupõe que, a despeito de suas diferenças étnicas e linguísticas, os índios compartilhem uma identidade comum, a partir da qual estabelecem entre si uma espécie de comunidade supraétnica. Essa imagem do índio genérico ao qual corresponde uma língua franca, entendida por todos, não deixa de ser um subproduto do processo colonizatório, com sua política de língua geral e da redução/conversão, como forma de apagamento das diferenças existentes entre as diversas nações indígenas.

E sobre a posição dos tupi e a presença de diversas outras etnias em território brasileiro, Mattoso Câmara Júnior (1977, p. 49) apresentou o mapa a seguir:



Mapa I: Mapa etnolinguístico da realidade indígena no Brasil segundo Mattoso Câmara Júnior (1977)

Nota-se que a realidade étnica brasileira em muito se distanciava da homogeneidade tupi apregoada por diversos relatos históricos. Também se deve atribuir essa “homogeneidade etnolinguística” à própria forma de conhecimento das realidades indígenas da época.

Tendo em vista que os índios não conheciam a escrita, sua forma de perpetuação cultural reservava-se apenas às artes de pintura corporal, sepultamentos e artesanato, ficando os relatos a cargo dos europeus, que descreviam a realidade de acordo com suas próprias concepções de mundo e suas experiências nessas terras. Porém, o fato de não dominarem a escrita não exclui o legado linguístico desses povos, uma vez que se falavam, à época, como já assinalado, diversas línguas indígenas.

Quanto à realidade linguística indígena desse momento histórico, por exemplo, Urban (1992, p. 87) argumenta que muito ainda está por se estudar:

[...] Estamos começando [no início da década de 1990] a conhecer suficientemente as línguas Tupi e nosso conhecimento das línguas Arawak está avançando bastante. Temos menos certeza quanto às línguas Jê e nosso conhecimento da família Karib é apenas rudimentar (URBAN, 1992, p. 87).

Assim sendo, se percorrido o território nacional a partir de seu extremo norte, teríamos este panorama: “a família etno-linguística pano ocupa, na região fronteira entre Brasil e Peru, uma área quase ininterrupta que se estende praticamente do Alto Solimões até o alto Purus” (ERIKSON, 1992, p. 239). A noroeste da Amazônia, segundo Wright (1992, p. 253), havia “três famílias linguísticas principais: Arawak, Tukano e Maku, com alguns grupos pequenos de Karib”.

No que diz respeito à atual região Nordeste brasileira, há predomínio da família Karib, que já registrava pequena presença no Norte, apresentando-se “desde o Ceará e a Paraíba até a porção setentrional do sertão baiano” (DANTAS; SAMPAIO; CARVALHO, 1992, p. 432). E quanto à família linguística, os estudos dos mesmos autores revelam que os limites da família Kariri são certamente extrapolados.

No restante do Brasil, grosso modo, divisavam-se dois grandes blocos que subdividiam toda a população indígena, ao longo de toda a costa e na bacia Paraná-Paraguai: “ao sul, os guarani, que ocupavam a bacia supracitada e o litoral, desde a lagoa dos Patos até a Cananéia, no atual estado de São Paulo; e os Tupi que dominavam a faixa litorânea desde Iguapé até, pelo menos, a costa do Ceará” (FAUSTO, 1992, p. 381-382).

Deve-se também ressaltar a presença marcante no Centro-Oeste brasileiro dos índios Payaguá e Guaikuru, especialmente, exímios canoeiros e cavaleiros, respectivamente, que em muito empreenderam resistência à presença do

colonizador nessa região, além dos Kadiwéu, do grupo dos Guaikuru, reconhecidos como artesãos pacíficos.

Sobre a realidade linguística, acrescentando-se às informações já apontadas, segue um esquema proposto por Rodrigues e adaptado por Melatti (RODRIGUES, 1975 *apud* MELATTI, 1980, p. 36-40), que apresenta as línguas indígenas brasileiras vivas, documentadas separadas em seus troncos e famílias⁷:

| TRONCO | FAMÍLIA | LÍNGUA |
|------------------------------|----------------|----------------------|
| Tupi | Tupi-Guarani | Guarani |
| | | Tenetehára |
| | | Apiaká |
| | | Tapirapé |
| | | Kamayurá |
| | | Kawahíb |
| | | Awetí |
| | | Mawé |
| | Mundurukú | Mundurukú |
| | Juruna | Juruna |
| | Arikê | Karitiâna |
| | Tuparí | Tuparí |
| | | Makuráp |
| | | Kepkiriwát |
| | Ramarâma | Itogapúk |
| | | Arara (de Rondônia) |
| | Mondé | Salamâi |
| | | Gavião (de Rondônia) |
| | | Suruí (de Rondônia) |
| | | Cinta Larga |
| Não classificadas em família | Masaká | |
| | Avá | |
| | Guajá | |

| TRONCO | FAMÍLIA | LÍNGUA |
|-----------------|----------------|---------------|
| Macro-Jê | Jê | Timbira |

⁷ Para a organização desse quadro, as línguas foram agrupadas por famílias, dada a similaridade de seus vocábulos e gramáticas, e essas famílias, cujas características se assemelhem, são agrupadas em troncos (antes também chamados blocos).

| | | |
|--|-------------------------------------|--------------------------|
| | | Kayapó |
| | | Suyá |
| | | Akuên |
| | | Kaingáng |
| | | Xokléng |
| | Maxakalí | Maxakalí |
| | | Pataxó (não mais falada) |
| | Karirí | Kirirí (não mais falada) |
| | Borôro | Borôro |
| | | Umutína |
| | Ainda não classificadas em famílias | Krenák (não mais falada) |
| | | Iatê (Fulniô) |
| | | Karajá |
| | | Guató |
| | | Eripaktsá |

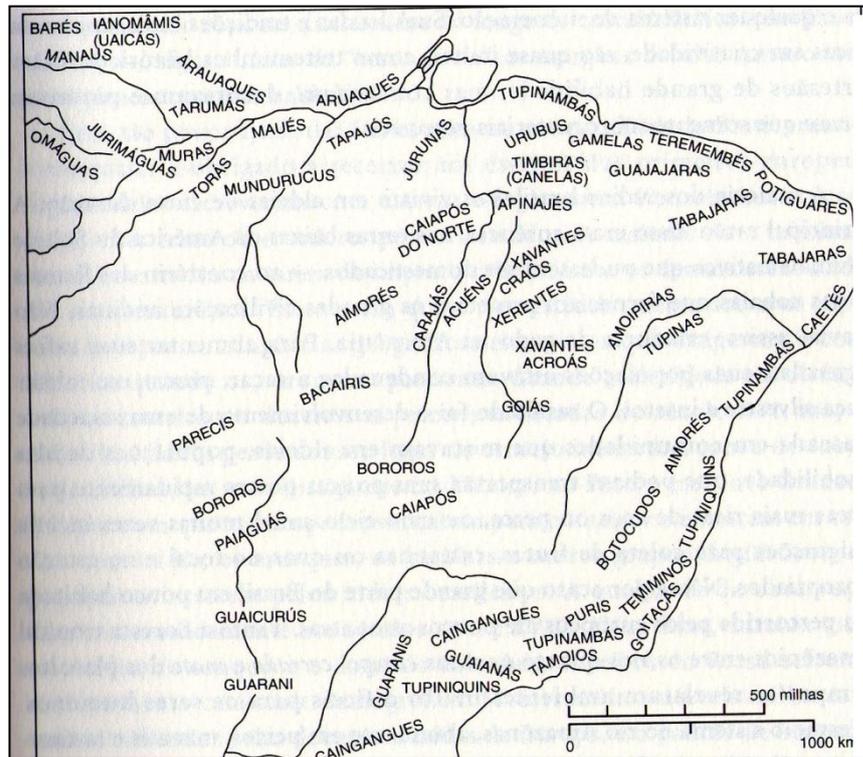
| <i>TRONCO</i> | <i>FAMÍLIA</i> | <i>LÍNGUA</i> |
|---------------|----------------|---------------|
| Aruák | Aruák | Guaná |
| | | Paresí |
| | | Manitenéri |
| | | Baniwa |
| | | Tariâna |
| | | Mandawáka |
| | | Wapitxâna |
| | | Palikúr |
| | Arawá | Kulina |
| | | Yamamadí |
| | | Daní |

| | <i>FAMÍLIA</i> | <i>LÍNGUA</i> |
|-----------------------------------|----------------|---------------|
| Ainda não classificadas em tronco | Karíb | Bakairí |
| | | Galibí |
| | | Apalaí |
| | | Mayongóng |
| | | Tiriyó |
| | | Taulipáng |
| | Makú | |
| | Yanoâma | Sanüamá |
| | | Yanâm |

| | | |
|------------|--|------------------------|
| | | Yanomámü |
| Tukâno | | Tukâno |
| | | Kobéwa |
| Katukína | | Katukína (do Amazonas) |
| | | Kanamari |
| Múra | | Múra-Pirahá |
| Pâno | | Kaxináwa |
| | | Nukiní |
| | | Poyanáwa |
| | | Amawáka |
| | | Yamináwa |
| | | Katukína (do Acre) |
| | | Marúbo |
| | | Mayorúna |
| | | Kulína |
| | | Karipúna (de Rondônia) |
| Txapakúra | | Pakaanóva |
| Nambikuára | | Nambikuára do Norte |
| | | Nambikuára do Sul |
| | | Sabanê |
| Guaikurú | | Kadiwéu |

| FAMÍLIA | LÍNGUA |
|------------------------------------|--------------------------|
| Ainda não classificadas em família | Tükúna |
| | Kanoê |
| | Irântxe e Menkü |
| | Trumái |
| | Xukurú (não mais falada) |

Já o mapa, a seguir, de etnias indígenas, para além do anteriormente apresentado, visualiza a diversidade dos povos indígenas em terras brasileiras. Proposto por Bethell em sua obra acerca da História da América Latina (BETHELL, 2004, p. 103), apresenta também diversas outras etnias que ocupavam outros espaços nacionais:



Mapa II: Etnias indígenas do Brasil segundo Bethell (2004)

Em síntese, pode-se depreender que, como assinalado anteriormente, a ideia de homogeneidade etnolinguística no Brasil durante a colonização não se confirma, mas poderia ser justificada pelo fato de os primeiros colonizadores terem conhecido inicialmente apenas os índios tupi, que dominavam a costa nacional, o que não representa, conforme o exposto, a real situação étnica daquele período.

Além, pois, do primeiro e principal contato dos portugueses com os índios tupi, outra situação que favoreceu o conhecimento das línguas dessa etnia foi o seu estudo pelos padres jesuítas. Um capítulo relevante na História do Brasil é a presença desses religiosos e sua respectiva influência na formação da sociedade colonial, na economia e mesmo na formação linguística do povo brasileiro. Isso porque a Companhia de Jesus aportou neste solo a partir de 1549, ainda no início do período colonial, com o intuito de catequizar os gentios na religião católica. Os jesuítas tinham o apoio da Metrópole em sua missão, uma vez que seriam mais uma frente para a colonização. Inicialmente, os religiosos estudaram a língua tupi para

utilizá-la em seu “processo de conversão”, ou seja, o conhecimento da língua funcionaria como elo de aproximação com os gentios.

Essa estratégia trouxe resultados altamente positivos para o processo de colonização, uma vez que grandes aldeamentos, as chamadas reduções, passaram a ser formados pelos jesuítas, juntando, por vezes, milhares de indígenas catequizados e, porque não dizer, “domesticados” na rotina europeia.

Então, com vistas a conhecer a língua indígena para auxiliar no trato com os autóctones, os estudos jesuíticos acerca da língua tupinambá renderam diversas produções religiosas na língua indígena, como sermões, ladainhas e mesmo produções literárias. Dentre eles, destacam-se as obras de um jesuíta aqui chegado ainda rapaz, com 19 anos, José de Anchieta, responsável por grande parte das obras de referência sobre essa língua tupi⁸, como, por exemplo, a *Gramática da língua mais usada na costa do Brasil* (1595), obra descritiva da língua tupinambá, do tronco Tupi. Todavia, essa descrição seguiria os padrões da gramática latina, em voga na época, o que significa que se procurava “enquadrar” a língua indígena aos padrões já conhecidos e estudados de línguas canônicas, falseando, por vezes, a real essência da língua tupi. E, dessa maneira, via catequizaç o, os indígenas também eram apresentados à língua portuguesa, essa sim digna de um estudo pormenorizado e refinado, segundo as orientações da Metr pole, tendo em vista que serviria como um portal de acesso à sociedade colonial. Segundo Mariani (2001, p. 110), a língua indígena

[...] serve como instrumento inicial de doutrinação, mas seu valor é anulado após o conhecimento da religião. Ou seja, do ponto de vista do rei e dos missionários, a língua indígena é útil apenas em um primeiro momento para que se possa capacitar o índio na religião católica. Porém, tão importante quanto conhecer a religião católica é saber a língua portuguesa, pois é através da língua portuguesa que

⁸ Nome genérico com que se designou a língua tupinambá, utilizada à época da colonização.

mais facilmente o índio receberá a fé. [...] Assim, tanto a língua portuguesa quanto a religião católica são necessárias sempre, enquanto que as línguas indígenas podem com o tempo vir a ser descartadas.

Mesmo assim, o contato entre o português e o índio, especialmente o tupi, deixou marcas profundas na formação da sociedade brasileira e, em decorrência disso, também em sua herança linguística. Num ambiente tão diverso do europeu, o colonizador aprendeu com o índio

[...] a identificar, a denominar e a classificar e usar toda a natureza tropical. [...] Aprenderam, igualmente, com eles, técnicas eficazes ajustadas às condições locais e às diferentes estações do ano, relativas ao cultivo e preparação de variados produtos de suas lavouras, à caça na mata e à pesca no mar (RIBEIRO, 1995, p. 129).

Ratificam essa afirmação os ciclos econômicos do tabaco, da lavoura de mocó e mesmo as técnicas de extração do pau-brasil e a criação de gado, por exemplo, atividades amplamente conhecidas e empreendidas pelos autóctones, que ensinaram seus conhecimentos aos brancos.

Dessa maneira, mesmo o índio tendo sido aculturado gradativamente pelos agentes colonizadores, deixou marcas linguísticas, econômicas e sociais na formação da sociedade brasileira, refletidas até hoje em hábitos, tradições e designativos ainda em uso. Além disso, decorrido esse primeiro momento de aculturação, os jesuítas passaram também a ver o indígena como o verdadeiro proprietário dessas terras, assim como a respeitá-lo em sua cultura, o que fez com que as reduções jesuítas passassem a defender os gentios contra os desmandos da Metrópole e mesmo dos colonos, ávidos por uma mão-de-obra barata e de fácil aquisição.

Em consequência disso, os jesuítas, com o tempo, começaram a ser combatidos pela Coroa Portuguesa, que via nas reduções, após um primeiro momento de paz, agora um entrave à unificação do território brasileiro, tendo em

vista que os aldeamentos tinham modo de viver próprio, com suas regras de comércio e produção. Isto é,

[...] obedeciam a regras próprias [...] e tinham uma política definida com relação a questões vitais da colonização, como a indígena. Além disso, na medida em que se tornaram proprietárias de grandes extensões de terra e empreendimentos agrícolas, as ordens religiosas não dependiam da Coroa para sua sobrevivência (FAUSTO, 2008, P. 61).

Além disso, pela profunda propagação da língua indígena, a Coroa Portuguesa também via com maus olhos o trabalho jesuítico, tendo em vista que, por utilizarem a língua nativa no trato inicial com os gentios, o não uso da língua portuguesa poderia resultar numa dificuldade ainda maior em unir o vasto território brasileiro. Tal situação culminou na expulsão desses religiosos do Brasil e mesmo de Portugal pela Coroa, em 1760, e a administração colonial passou a contar mais com a nova população formada, especialmente por mestiços, para sedimentar a identidade nacional, que receberia depois mais uma etnia que lhe forneceu bases, os negros, que vieram trabalhar na extração e na manufatura da cana-de-açúcar.

1.1.1.1 O ciclo⁹ da cana-de-açúcar e a estruturação da economia colonial

Como referido no tópico anterior, nota-se uma identificação do índio com os processos de extração, caça e pesca na Colônia, pelo menos de acordo com o olhar europeu. Dessa forma, a atividade que foi tomando força ainda no final do século XVI, substituindo gradativamente a atividade meramente extrativista do pau-

⁹ O historiador Boris Fausto (2008) discorda da divisão da economia colonial em ciclos econômicos no Brasil Colônia, haja vista que essa denominação iria contra a própria história, uma vez que ciclos representam um início, meio e fim de uma etapa para só depois se iniciar uma próxima, o que, de fato, não ocorre nos registros das atividades econômicas brasileiras no Período Colonial. Contudo, aqui a utilizamos por já estar cristalizada nos relatos históricos e também por assim vigorar em diversas das obras que serviram de base para este texto.

brasil, foi a lavoura da cana-de-açúcar, e que precisaria de uma mão de obra mais submissa e disciplinável, a escrava, para além da indígena.

E, considerando que a África tinha contingente humano substancial à época e o comércio escravista já era uma realidade na Europa, os escravos africanos foram importados para o Brasil. Havia também, por parte dos senhores de engenho, certas preferências em relação à origem desses negros, o que, segundo eles, identificaria uma maior ou menor adequação a certos tipos de trabalho. Segundo Antonil (1997¹⁰, p. 89),

[...] Os que vêm para o Brasil são ardas, minas, congos, de São Tomé, de Angola, de Cabo Verde e alguns de Moçambique. [...] Os ardas e os minas são robustos. Os de Cabo Verde e de São Tomé são mais fracos. Os de Angola, criados em Luanda, são mais capazes de aprender ofícios mecânicos. [...] Entre os congos, há também alguns bastantemente industriais e bons não somente para o serviço de cana, mas para as oficinas e para o meneio da casa.

Assim, com o tráfico negreiro autorizado a partir de um alvará de D. João III, em 1549 (LUCCHESI, BAXTER, RIBEIRO, 2009, p. 45), a importação de negros cresceu de tal forma que os africanos passaram a ser maioria na sociedade colonial. E esse comércio era cercado de vantagens: além de atender a um público ávido por mão de obra e que tinha condições de pagar um bom preço pelas ‘peças’ – nome pelo qual eram designados os negros produtos da escravidão –, rendia também lucros inimagináveis para seus vendedores, uma vez que, a princípio, a captura de africanos era relativamente fácil, isso pelo auxílio dos próprios africanos na captura de escravos, por meio de argumentos sedutores de melhores condições de vida e pelo imenso contingente de negros naquele continente.

No Brasil, eles ocupavam todos os postos na cultura de cana-de-açúcar. Desde a extração até a lida com o produto final e, como relatado por Antonil, alguns

¹⁰ A primeira edição da obra data de 1711.

foram também designados para os lides do lar, nas chamadas Casas Grandes. A essa época, com o despontar da produção de açúcar não apenas no Brasil, mas também nas Antilhas, promovida por holandeses, posteriormente, esse produto passou de especiaria apenas consumida comedido pelas classes europeias mais abastadas a produto de consumo diário, tal qual o conhecemos na atualidade.

No que diz respeito às línguas africanas faladas por esses povos, há que se compreender que, sendo eles de origens tão distintas, também suas línguas eram variadas. Então, “levando-se em conta a região de origem dos escravos africanos, pode-se estimar que foram atingidos pelo tráfico locutores de cerca de 200 a 300 línguas” (PETTER, 2001, p. 223). Assim, nem no Brasil havia um monolinguismo, nem nas nações africanas de onde partiram seus habitantes para essas terras; o plurilinguismo era factual, com um grande número de línguas indígenas que coexistiam no país, assim como as diversas línguas dos povos africanos que aqui aportaram na condição de escravos.

Com isso, fazia-se necessário o uso de uma língua para a comunicação entre negros e entre negros e seus senhores, papel ocupado pela língua portuguesa, já com inserções indígenas, mas especialmente transmitida pelos índios chefes da produção açucareira e depois pelos negros ladinos, que a aprendiam, usavam-na no trato com seus senhores e também na comunicação com os demais moradores da senzala.

É necessário ainda que se considere que, apesar de ter como base o negro, também o índio, ainda que em menor proporção, trabalhava nos engenhos de cana, por vezes também como escravo, uma vez que o comércio de indígenas praticado especialmente pelos paulistas também se tornou lucrativo,

Além disso, também nesse período, e com vistas a coibir a fácil entrada de outras nações no vasto território brasileiro, a Coroa Portuguesa pôs em execução o regime das donatarias, que formariam as Capitanias hereditárias. O sistema consistia na doação de grandes porções de terras a grandes senhores, que seriam responsáveis pelo povoamento e pela administração da porção de terra a ele pertencente. Em outras palavras, “o donatário era um grã-senhor investido de poderes feudais pelo rei para governar sua gleba”, isto é, “com o poder político de fundar vilas, conceder sesmarias, licenciar artesãos e comerciantes, e o poder econômico de explorar diretamente ou através de intermediários suas terras e até com o direito de impor a pena capital” (RIBEIRO, 1995, p. 87).

Porém, logo esse regime administrativo deu mostras de sua ineficiência: apenas algumas capitanias, como São Vicente e Pernambuco, por exemplo, alcançaram êxito em suas atividades econômicas e de gestão. Nesse cenário, a Coroa Portuguesa instituiu, então, o Governo Geral, para auxiliar a governança das donatarias e também com vistas a melhor defender o território.

E, tendo em vista as riquezas naturais da época, como a cana-de-açúcar, por exemplo, e sua concentração no Nordeste, o rei D. João III instituiu Tomé de Sousa como primeiro Governador-Geral e a sede do Governo na Bahia de Todos os Santos.

E, com sua base no trabalho escravo negro, a cana-de-açúcar foi o ciclo mais duradouro na economia colonial, iniciando-se ainda em meados do século XVI e permanecendo até o século XVIII. Essa atividade estava também ainda concentrada em áreas litorâneas, especialmente no Nordeste.

Assim,

[...] O engenho açucareiro, primeira forma de grande empresa agroindustrial exportadora, foi, a um tempo, o instrumento de

viabilização do empreendimento colonial português e a matriz do primeiro modo de ser dos brasileiros. Sem ele, [...] seria inimaginável a ocupação europeia de uma vasta área tropical, sem riquezas minerais para descobrir, habitada por indígenas que apenas lograram construir culturas agrícolas e que não constituíam uma força de trabalho facilmente disciplinável e explorável (RIBEIRO, 1995, p. 274).

E o mercado consumidor do açúcar brasileiro foi também o grande responsável por esse sucesso econômico colonial. Tendo se tornado produto de uso diário, à medida que todas as mesas utilizavam-se do açúcar em substituição ao mel e, dada a experiência anterior dos portugueses em seu cultivo, o Brasil dominou esse mercado por muito tempo.

Essa hegemonia foi ameaçada apenas a partir das invasões holandesas no Nordeste brasileiro entre 1624 a 1654, visando à posse dos lucros advindos desse comércio. As invasões foram combatidas e vencidas, mas a experiência dos holandeses com o açúcar os levou a cultivar a cana em outras terras, como as Antilhas, que concorreram diretamente com o açúcar brasileiro e forçaram a Colônia Portuguesa a procurar outras fontes de renda.

1.1.1.2. O ciclo do ouro e a alteração de polos no Brasil Colônia

Como visto, a Capitania de São Vicente se destacou na época colonial, seja pelas fundações proporcionadas pela Coroa ainda antes de se tornar donataria, seja pela formação de um engenho de açúcar. Todavia, essa porção sul estabeleceu-se como polo econômico e cultural, de fato, a partir do descobrimento “do ouro de lavagem na Capitania e a crescente importância das entradas de caça ao índio, [...] tornando-se São Vicente a grande porta de entrada do Sul, do Oeste e do Centro da colônia, ainda desconhecidos” (VIANNA, 1965a, p. 75).

Ao se referir a essa porção da sociedade brasileira colonial, Ribeiro (1995, p. 364) dá-lhe o nome de Brasil caipira, caracterizando a região como importadora de escravos e centro de mineração. Contudo, para se entender como São Vicente chegou a esse patamar, devem-se elencar diversos fatores e acontecimentos históricos que culminaram, inclusive, na mudança da capital do Governo para o Rio de Janeiro (1763).

A Colônia Espanhola já havia descoberto ouro e outros minérios na América do Sul, o que lhe aumentou o poderio no cenário mundial e a tornou mais ameaçadora ao domínio português, com invasões cada vez mais frequentes ao atual Sudeste do Brasil. Com isso, rumores de que ainda na porção portuguesa haveria esse tipo de riquezas fizeram com que a busca por minerais acabasse também por promover a expansão do território.

Como já referido anteriormente, sabe-se que a sociedade da época tinha como base de sua constituição a miscigenação entre brancos e índios, inicialmente, e que depois recebeu contribuições dos negros africanos. Caracterizada então como uma população híbrida, havia polos em que o negro prevalecia e outros em que o sangue indígena dominava. Segundo Prado Júnior (1986, p. 112),

[...] O preto e seus derivados dominam nas regiões de grande atividade econômica. [...] É assim que ele se concentra nos grandes núcleos agrícolas do litoral: no Maranhão, no Extremo-Nordeste, no recôncavo baiano, no Rio de Janeiro, assim também nos centros mineradores do interior: Minas Gerais, Goiás, Mato Grosso.

Já o índio predominava nas áreas que iam da Amazônia até a porção Centro-Oeste do Brasil, “em geral as de vida econômica pouco ativa” (PRADO JR, 1986, p. 113). Além disso, a mistura de raças tinha sempre o branco numa posição inicial, seja branco e índio ou branco e negro. As uniões com os índios, como já verificado, foram até incentivadas pela Metrópole por algum tempo, enquanto as

relações entre brancos e negros davam-se especialmente de forma ilícita, entre senhores e escravas, que geravam filhos bastardos e não reconhecidos. Já os índios que se relacionavam com negros eram considerados como “de sangue sujo”, declinando na escala social imposta pela sociedade da época.

E o Sudeste caracterizou-se principalmente pela população de mamelucos, isto é, mestiços de brancos e índios, a quem Darcy Ribeiro também dá o nome de *brasilíndios*. Ainda segundo o antropólogo,

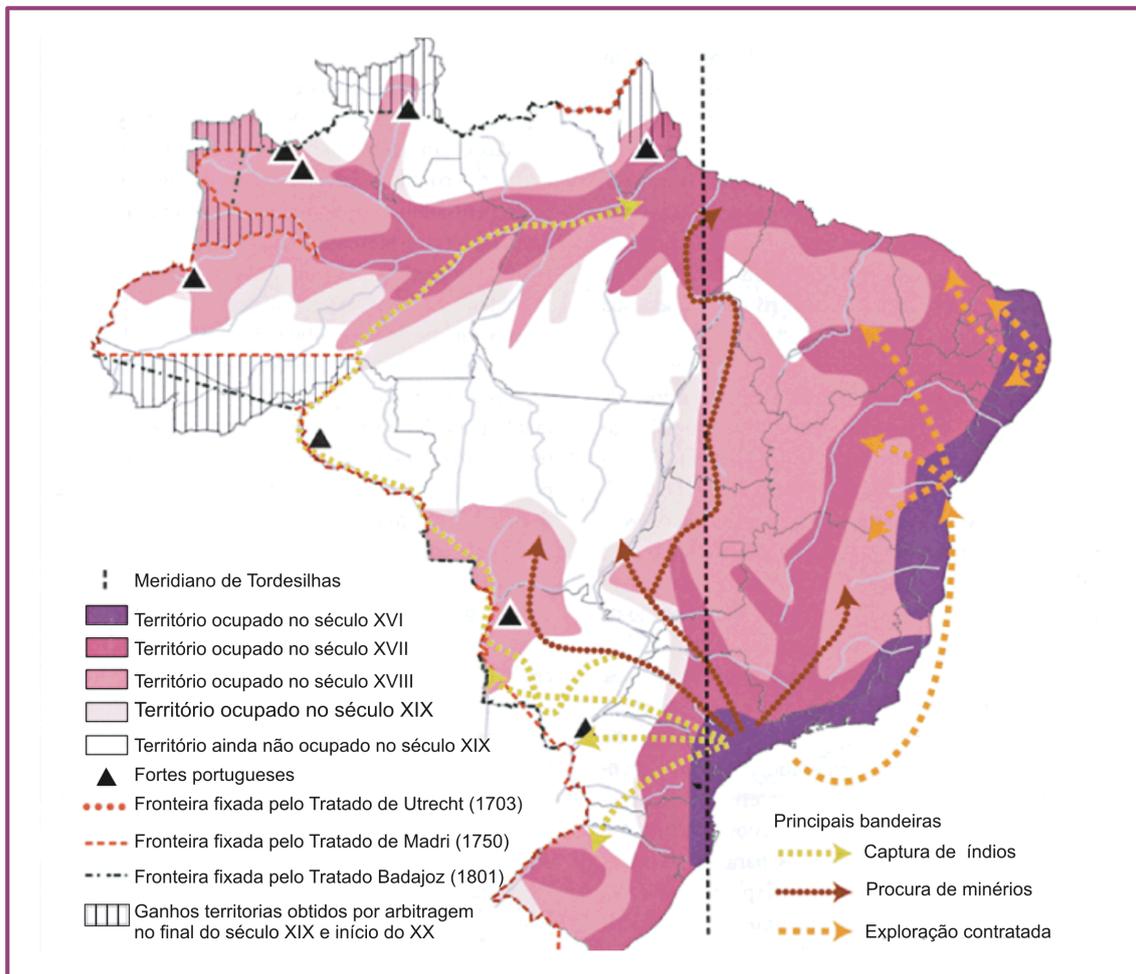
[...] a expansão do domínio português terra adentro, na constituição do Brasil, é obra de *brasilíndios* ou mamelucos, [...] [que] dilataram o domínio português exorbitando a dação de papel das Tordesilhas, excedendo a tudo que se podia esperar (RIBEIRO, 1995, p. 106).

Isso ocorreu em virtude de os índios, já conhecedores das terras por meio da herança que receberam de seus ancestrais, também terem auxiliado os portugueses a adentrar na mata e foram base para a formação das entradas e bandeiras, que redesenharam o território brasileiro. “O movimento *entradista* partiu no século XVI de vários pontos da Costa Leste para o interior ainda desconhecido, [...] sendo seu último capítulo a conquista do Acre, já no Brasil imperial e republicano” (VIANNA, 1965a, p. 192). Já

[...] O *bandeirantismo*, designação dada ao movimento de penetração realizado principalmente por moradores da Capitania de São Vicente (depois denominada de São Paulo), começou com o *ciclo da caça ao índio*, ainda na segunda metade do século XVI; entrou no setor da mineração com o *ciclo do ouro de lavagem*; assumiu, depois, o original aspecto apresentado pelo *ciclo do sertanismo de contrato*; continuou já em fins do século XVII e primeira metade do seguinte, com o *grande ciclo do ouro* e terminou, ainda na era setecentos, nos *ciclos de povoamento* (VIANNA, 1965a, p. 192).

O mapa apresentado a seguir (NELI; MELLO, 2005) resume os principais caminhos tomados pelas bandeiras em território por vezes brasileiro por vezes ainda de domínio espanhol. Além disso, demonstra claramente a expansão do território, ao

ilustrar o Meridiano de Tordesilhas¹¹ e o posterior Tratado de Madri¹², bem como mostra, por meio de cores, a ocupação do território durante os séculos XVI a XX:



Mapa III – Rotas das bandeiras e seus objetivos – séculos XVI a XIX (NELI; MELLO, 2005)

Assim, a capitania de São Vicente, principal origem dos movimentos de expansão territorial, era notadamente constituída por mamelucos, isto é, com notável

¹¹ Pelo Tratado de Tordesilhas (1494), intermediado por Alexandre VI, estabeleceu-se um meridiano a 370 léguas a oeste de Cabo Verde e, a partir dele, instituiu-se que os espanhóis teriam direito à exploração das terras a oeste e os portugueses, a leste (Brasil Escola, disponível em <http://www.brasilecola.com/historiab/tratado-de-tordesilhas.htm>. Acesso em: 12 jan 2012).

¹² O Tratado de Madrid foi firmado entre D. João V, de Portugal, e D. Fernando VI de Espanha, e substituiu o Tratado de Tordesilhas. Por meio dele, definiu-se que valeria o princípio do direito privado romano do *uti possidetis, ita possideatis*, que estabelece que as terras colonizadas de fato fossem também de direito de seus colonizadores, independente do Meridiano de Tordesilhas. Dessa forma, delimitaram-se então os contornos aproximados do Brasil atual (InfoEscola, disponível em <http://www.infoescola.com/historia/tratado-de-madrid-de-1750/>. Acesso em 12 jan 2012).

influência indígena. Segundo Fausto (2008, p. 93), “os portugueses de São Paulo adotaram muitos dos hábitos e habilidades indígenas, tornando-se tão capazes de usar o arco e a flecha como as armas de fogo”. Isso sem considerar a língua dominante na donataria, a Língua Geral, também reflexo dessa influência indígena, especialmente a tupi.

Então as bandeiras¹³ paulistas tornaram-se instrumento de Portugal, mesmo que por vezes indiretamente, haja vista que algumas eram financiadas com recursos particulares, para a expansão do território e para a efetiva ocupação e posse das terras brasileiras.

Como as bandeiras também tinham como propósito o aprisionamento de índios e eram formadas em grande parte por mamelucos, os descendentes dos índios passaram a caçá-los nas tribos para o trabalho escravo, seja como uma alternativa ao alto custo dos negros africanos, seja como uma opção para outros trabalhos com os quais mais se identificavam, como a lida com o gado e o trabalho na lavoura. Além disso, as bandeiras partiram de São Vicente também em busca de metais preciosos, como se referiu Vianna (1965a). De acordo com Prado Jr, as minas foram “sucessivamente descobertas a partir dos últimos anos do século XVII, [fixando] núcleos estáveis e definitivos no coração do continente (Minas Gerais, Goiás, Mato Grosso)” (PRADO JR, 1986, p. 37), contribuindo, pois, tanto para o aumento das riquezas do país quanto para o povoamento do território.

Com a descoberta dos minérios, o interesse da Metrópole voltou-se para essa porção secundária da Colônia, uma vez que, até o início do século XVIII, ainda era a cana-de-açúcar que dominava o cenário econômico do Brasil. Com a

¹³ Bandeirantismo foi o nome atribuído ao movimento de penetração dos moradores da Capitania de São Vicente, iniciado na segunda metade século XVI (ciclo da caça ao índio) e que se fortaleceu no final do século XVII, com o descobrimento do ouro (ciclo do ouro de lavagem), concluindo-se nesse mesmo século com ciclos de povoamentos (VIANNA, 1965b, p. 192).

instituição das minas no Sudeste e conseqüente concentração de riquezas para seus governantes, logo a sede administrativa do Vice-Reinado passou de Salvador para o Rio de Janeiro (1763), ponto por onde entravam escravos e suprimentos e também por onde saía o ouro das minas.

Desenhou-se, pois, um território muito próximo ao Brasil atual, destacando-se especialmente o papel das entradas e bandeiras para o povoamento das áreas descobertas. Além disso, deve-se creditar a essas expedições também a expansão da Língua Geral no Brasil Colonial a partir de sua disseminação pelos expedicionários, tanto no uso corrente quanto na atribuição de topônimos de base indígena pelos lugares por onde passavam. Por eles,

[...] O tupi [base linguística da Língua Geral] se espalha, historicamente, do Espírito Santo, por São Paulo e Minas Gerais e, subindo, por grande parte de Goiás e Maranhão, até a embocadura do Amazonas – numa descrição atual, isto é, posterior aos contatos demográficos dos séculos XVI e XVIII (HOUAISS, 1985, p. 53).

Também o ciclo do gado foi resultado dessa expansão, inicialmente porque era necessário um meio de transporte para o ouro, mas também porque era o destino de boa parte dos índios capturados e vendidos como escravos pelas bandeiras de preação¹⁴. Vianna (1965a, p. 230) esclarece que a criação de gado proporcionou “a ligação geográfica dos movimentos de expansão partidos da Bahia e de São Vicente, de Pernambuco e do Maranhão”. Assim, “unidos, no Norte de Minas, no primeiro caso, no interior do Piauí ou do Ceará, no segundo, por intermédio dos passadores de gado, processou-se a verdadeira união terrestre do Sul, Centro e Nordeste” (VIANNA, 1965a, p. 230).

¹⁴ Aquelas que tinham como objetivo a captura de indígenas.

1.1.1.3 A crise no sistema colonial

A relativa independência dos bandeirantes frente ao governo português, todavia, por vezes trazia também problemas para a administração portuguesa das terras além-mar. Além disso, as dimensões continentais do território brasileiro eram difíceis de gerir, especialmente por Portugal, que enfrentava uma séria crise econômica nos idos dos séculos XVIII e XIX.

Associado a essas condições, elenca-se também o cenário mundial da época, especialmente no que diz respeito a essa forma de expansão comercial, o colonialismo: colônias inglesas e francesas tornaram-se independentes enquanto as monarquias absolutas sofriam rejeição com o advento das ideias iluministas e seu liberalismo defendido. A escravidão, base do colonialismo, era banida da França (1794) e da Inglaterra (1807) e passou a ser combatida também no Brasil por uma nova elite formada a partir das riquezas advindas especialmente do ouro e do gado (FAUSTO, 1995).

Portugal ainda tentou combater, em sua maior colônia, esses acontecimentos, que acabariam por arruinar o sistema colonial, sua maior fonte de renda à época. Prova disso foi a atuação do Marquês de Pombal, ministro de Dom José I, que expulsou os jesuítas do Brasil, integrou os índios definitivamente na sociedade colonial, reconhecendo-os como participantes dessa sociedade, “sujeitos dotados de alma”, e passou a controlar avidamente os lucros advindos da mineração.

Porém, em resposta a essa situação, começaram a surgir no Brasil movimentos de rebeldia, que visavam a combater a política portuguesa e a lutar pela

independência, dentre eles, a Inconfidência Mineira (1789), a Conjuração dos Alfaiates (1798) e a Revolução de 1817, em Pernambuco.

Aliando-se o intento de combater os movimentos pela independência à necessidade urgente de Portugal defender-se contra os ataques napoleônicos, a Corte Portuguesa desembarca no Brasil em 1808, transferindo a sede da monarquia portuguesa para sua Colônia e alterando o quadro das relações internacionais no contexto da América do Sul. Ainda registrou-se uma mudança visível no campo cultural, tendo em vista que a corte acabava por necessitar de educação e cultura, condições estendidas às elites coloniais.

Contudo, apesar dos esforços portugueses para manter o controle da colônia além-mar, as condições internacionais e mesmo as internas forçaram a Independência do Brasil, representada simbolicamente pelo Grito do Ipiranga dado por Dom Pedro I às margens do riacho com esse nome em 1822.

1.1.2 Período pós-colonial até a República: o estabelecimento da soberania nacional

A independência foi se consolidando paulatinamente no Brasil, inclusive com conflitos, especialmente militares.

Após 1822, com a Independência, tornou-se urgente a aprovação de uma Constituição, com a convocação de uma assembleia constituinte logo depois dissolvida por Dom Pedro, com o apoio dos militares: “a primeira Constituição brasileira nascia [assim] de cima para baixo, imposta pelo rei ao ‘povo’” (FAUSTO, 1995, p. 149). Leia-se aqui “povo” como a minoria de brancos e mestiços com direito a voto, uma vez que parte considerável da população era formada por escravos,

excluídos da vida social e, portanto, dos dispositivos constituintes, apesar do grande contingente desses escravos à época no Brasil.

Dessa maneira, tendo em vista que o maior contingente da população da época era composto por negros escravos, mulatos e índios, tribais ou mestiços, já em proporção semelhante a dos brancos europeus, grande parte dessa massa populacional era excluída das leis, o que pode ser exemplificado pela abolição da escravatura no Brasil definitivamente apenas em 1888, quase cem anos após o fim da escravidão na Europa.

A monarquia prevaleceu no Brasil mesmo depois da abdicação de Dom Pedro I (1831), após pressões internas e externas, com o período chamado de Regência (de 1831 a 1840), em que políticos governaram em nome do imperador, Dom Pedro II, até sua maioridade antecipada. E mesmo depois disso o regime monárquico perdurou, findando apenas em 1889, com a Proclamação da República.

Esses períodos foram marcados por revoltas em várias partes do Brasil, incentivadas pelas ideias liberais, em voga entre a sociedade letrada, e mesmo pelo poderio militar. “Enquanto as rebeliões agitavam o país, as tendências políticas no centro dirigente iam-se definindo” (FAUSTO, 1995, p. 171), com a formação de dois partidos, o Conservador, que reunia magistrados, burocratas, proprietários rurais e grandes comerciantes; e o Liberal, que agrupava a classe média urbana, alguns padres e proprietários rurais menos tradicionais.

Com isso, podemos identificar uma aproximação do Partido Conservador com a burocracia e o Liberal com os interesses agrários, opostos ao reforço do poder central.

E a elite agrária ganhou representatividade com o primeiro ciclo econômico iniciado no século XIX, a cultura do café, especialmente no Oeste Paulista.

1.1.2.1 O ciclo do café e a extração ervateira: um novo panorama étnico no Brasil

O café foi o propulsor da economia brasileira a partir de 1800, formando uma nova elite brasileira e alterando novamente o polo econômico, agora voltado para os paulistas. Inicialmente produzido no Vale do Paraíba (entre o Rio de Janeiro e São Paulo), e depois introduzido no interior paulista e no Paraná, esse grão foi o principal produto de exportação brasileiro por quase cem anos.

Além da importância econômica, o café também influenciou a sociedade da época, tendo em vista que mão de obra agora assalariada e qualificada no trato do grão passava a ser requisitada. Isso fez com que outros europeus e mesmo asiáticos aportassem no Brasil, em busca de melhores condições de vida, o que desenhou um novo panorama na realidade brasileira.

Outra atividade desenvolvida durante o mesmo período foi a extração da *Ilex paraguariensis*, a erva-mate, que, embora de menor importância comercial e econômica que o café, tem relevância para este estudo por congregar mão de obra indígena em sua extração, como os índios Guarani (TAVARES, 2005).

Ao final da Guerra da Tríplice Aliança, também conhecida como Guerra do Paraguai (1864-1870), o funcionário do governo brasileiro Tomás Laranjeira, ao se deparar com a existência de grandes áreas dotadas de extensos ervais, requereu para si a posse dessas terras, nascendo então a Companhia Erva-Mate Laranjeira,

que monopolizou o cultivo e a extração da erva-mate, de 1882 a 1937 em território sul-mato-grossense.

Essa atividade contava com o trabalho dos índios guarani e também de grande volume de mão de obra paraguaia, contribuindo, assim,

[...] diretamente para a formação étnica da população local, e para o povoamento da fronteira do Brasil com o Paraguai, visto que, com a chegada de muitos trabalhadores para executarem trabalhos nos ervais começaram a surgir novas povoações (TAVARES, 2005, p. 55).

Os guarani foram os primeiros a fazerem uso da erva-mate, por isso sua importância como trabalhadores nessa atividade econômica. Todavia, além da região Centro-Oeste, especialmente o Sul se beneficiou com a extração da erva-mate, esta se registrando, inclusive, como a atividade mais importante da Província Del Guairá da metade do século XVI até 1632, território que abrangia praticamente todo o Paraná.

Outros ciclos e entre ciclos econômicos também ocorreram concomitantes a esses, como os da borracha (1866-1913), da soja (a partir de 1970) e da indústria (a partir das primeiras fábricas trazidas ainda em 1808). Porém, como o intuito deste capítulo é introduzir o tema focalizado nesta dissertação e situá-lo historicamente, foram apontados aspectos socioeconômicos mais relacionados à inter-relação entre brancos e índios, e se entende que maiores detalhes desses momentos econômicos profundamente relevantes para o Brasil devem ser buscados em obras de referência acerca da História do Brasil, muitas delas que serviram, inclusive, de base para este estudo.

Assim sendo, o capítulo que trata brevemente da História do Brasil em seus aspectos sócio-político-econômicos encerra-se para dar lugar a aspectos

pontuais da pesquisa, como Aporte Teórico, Capítulo II; Metodologia, Capítulo III e Análise de dados, Capítulo IV; que serão tratados a seguir.

CAPÍTULO II

PRESSUPOSTOS TEÓRICOS: O LÉXICO EM FOCO

O homem é um ser social. Desde o momento em que nasce, ele se relaciona com seus iguais, expressa conhecimentos e ideias, além de perpetuar sua existência por meio da transmissão de crenças e de tradições para os indivíduos de seu meio social e interagir com outras comunidades culturais e linguísticas. Para esses contatos, ele se vale de um instrumento fundamental, a língua, em particular do léxico dessa língua, para a nomeação de elementos, objetos, situações e hábitos.

E essa relação entre língua e realidade social foi especialmente tratada a partir do início do século XX, inicialmente nos Estados Unidos, por Edward Sapir e, posteriormente, em associação com Benjamin Lee Whorf, por meio da teoria conhecida como Hipótese Sapir-Whorf, que concebe a língua como reflexo da visão de mundo, por meio da qual ideias são formadas, condicionadas.

As bases do Relativismo Linguístico, isto é, teoria para quem a língua estaria condicionada a seus falantes e ao universo em que se circunscreve, foi lançada por Edward Sapir e serve de pressuposto para este trabalho, tendo em vista sua vertente etnolinguística. Depois, seu discípulo teórico, Benjamin Lee Whorf, levou essa teoria ao extremo ao estudar as línguas indígenas americanas. Assim, os estudos de Sapir associados às contribuições de Whorf deram origem à Hipótese Sapir-Whorf, vertente teórica para a qual “todo sistema lingüístico manifesta, tanto no seu léxico como na sua gramática, uma classificação e uma ordenação dos dados da realidade que são típicas dessa língua e da cultura com que ela se conjuga” (BIDERMAN, 1998, p.93).

Segundo a mesma autora,

[...] O léxico de uma língua constitui uma forma de registrar o conhecimento do universo. Ao dar nomes aos referentes, o homem os classifica simultaneamente. Assim, a nomeação da realidade pode ser considerada como a etapa primeira no percurso científico do espírito humano de conhecimento do universo. Ao identificar semelhanças e, inversamente, discriminar os traços distintivos que individualizam esses referentes em entidades distintas, o homem foi estruturando o mundo que o cerca, rotulando essas entidades discriminadas. É esse processo de nomeação que gerou e gera o léxico das línguas naturais (BIDERMAN, 2006, p. 35).

Assim, no léxico, a representação de realidade ocorre porque dado elemento cultural passa a ser conhecido e sua existência compartilhada entre os indivíduos a partir do momento em que é nomeado. É como que passar para o campo do real dependesse, e de fato não prescinde disso, de ser designado, para que a vida de um elemento cultural seja partilhada por um grupo. Nesse sentido, o nível lexical da língua media também a relação entre língua e sociedade, língua e cultura, tornando-se, pois, um patrimônio cultural, sendo assim não apenas um produto, mas também elemento participante dessa mesma cultura, aqui entendida como todo o conhecimento de um grupo, que o distingue dos demais¹⁵.

Desta feita, concebe-se o léxico como “o nível da língua que melhor documenta o modo como um povo vê e representa a realidade em que vive” e entende-se “que o vocabulário de um grupo social atesta seus valores, suas crenças e também a forma como nomeia os referentes do mundo físico e do universo cultural em diferentes épocas da sua história” (ISQUERDO, 2003, p. 165).

Por registrar o conhecimento de um grupo e representar seu universo, o léxico reflete, assim, traços característicos da cultura na qual está inserido, expressando aspectos da identidade social de um grupo de indivíduos que perpassa

¹⁵ Entende-se cultura em sua rubrica antropológica, segundo a qual se trata de “conjunto de padrões de comportamento, crenças, conhecimentos, costumes etc. que distinguem um grupo social” (HOUAISS, 2001).

os limites linguísticos, além de ser influenciado também por aspectos sociais, geográficos, econômicos e históricos, por exemplo. Assim, de acordo com o grupo que o utiliza, esse léxico vai se desenhando com cores específicas, particulares, por vezes se assemelhando ao léxico geral de um grande grupo de usuários e por vezes se particularizando, com especificidades da realidade sociocultural de pequenos conglomerados humanos, caracterizando-se, assim, como a norma lexical desse grupo.

Por essas razões, o léxico pode ser considerado, segundo alguns teóricos, o nível linguístico que mais se aproxima da realidade cultural de um grupo, e seu estudo pode demonstrar características para além dos traços linguísticos de uma sociedade. Nesse sentido, a particularização do léxico também se deve a fatores relacionados à interação interétnica, pois, se duas comunidades linguísticas distintas interagem num mesmo ambiente, também seu conhecimento lexical é partilhado, uma vez que “é fenômeno comum, nas zonas de contato, a troca de palavras ou de expressões que acompanham o deslocamento das populações e dos objetos culturais a que se aderem” (DICK, 2006, p. 05). E é assim, caracterizando e particularizando o ambiente em que se insere e representando o pensar de um grupo social e suas inter-relações, que se entende o léxico no âmbito deste trabalho.

2.1. O CONCEITO DE NORMA LEXICAL E SUA APLICABILIDADE PARA O ESTUDO DE INDIGENISMOS NO LÉXICO DO PORTUGUÊS DO BRASIL

Os usuários da língua têm a sua disposição todo o sistema linguístico, com suas possibilidades de realização abertas para as escolhas individuais e coletivas dos falantes.

O dinamarquês Hjelmslev, em seu trabalho *Langue et parole*, de 1943, abordou essa questão das escolhas individuais e coletivas e as sistematizou numa tríade a que chamou esquema, norma e uso. Retomando a dicotomia de Saussure (1916), associou seu esquema à língua e o uso à fala, na visão saussuriana. A norma, para ele, seria “*forma material, definida por una realización social dada, pero independientemente todavía del detalle de la manifestación*”¹⁶ (HJELMSLEV, 1943 *apud* LARA, 1976, p.38), enquanto o uso representaria o “*conjunto de hábitos adoptados por una sociedad determinada y definidos por las manifestaciones observadas*”¹⁷ (HJELMSLEV, 1943 *apud* LARA, 1976, p.38).

Assim, sintetizando o conceito de norma hjelmsleviano, pode-se entender que “*la norma determina (es decir, presupone) el uso y el ato, y no inversamente. [...] el ato y el uso preceden lógicamente y prácticamente a la norma; la norma nace del uso y del ato y no al contrario*”¹⁸ (HJELMSLEV, 1943 *apud* LARA, 1976, p.43).

Os conceitos de sistema e de norma associados, por sua vez, foram sistematizados primeiramente por Eugênio Coseriu (1979), linguista estruturalista que, a partir também da dicotomia saussuriana *langue* e *parole* (língua e fala), desenvolveu a tricotomia “sistema, norma e fala”, para demonstrar que, para além da língua e da fala, existe um nível intermediário, o da norma, já referida por Hjelmslev (1943), que representa a realização do sistema propriamente dita.

Assim, o conceito coseriano de sistema estaria para o saussuriano de língua, e a concepção de fala seria a mesma para ambos. Todavia, Coseriu inclui ainda um nível intermediário, a norma, que

¹⁶ “A forma material, definida por uma realização social dada, mas independente, porém do detalhe da manifestação” (TN = tradução nossa).

¹⁷ “Conjunto de hábitos adotados por uma sociedade determinada e definidos pelas manifestações observadas” (TN).

¹⁸ “A norma determina (isto é, pressupõe) o uso e o ato, e não inversamente. [...] O ato e o uso precedem lógicamente e praticamente a norma; a norma nasce do uso e do ato e não o contrário” (TN).

[...] é, em certo sentido, mais ampla que o sistema: com efeito, ela encerra também os traços não funcionais, enquanto o sistema contém só os traços distintivos necessários para que uma unidade da língua (quer no plano de expressão, quer no do conteúdo), não se confunda com outra (COSERIU, 1980, p. 123).

O linguista romeno, para além de conceber o conceito de norma e discorrer sobre as relações entre norma e sistema, também estabelece:

[...] Em outro sentido, o sistema é mais amplo que a norma, já que é menos determinado. [...] De fato, o sistema, contendo apenas as oposições funcionais, encerra também tudo o que na língua seria possível, mesmo se não é realizado na norma. Em outras palavras, a norma abrange fatos linguísticos efetivamente realizados e existentes na tradição ao passo que o sistema é uma *técnica aberta* que abrange virtualmente os fatos ainda não realizados (COSERIU, 1980, p. 123).

Então, se compararmos as tríades de Hjelmslev e de Coseriu, notamos que o *esquema* do primeiro está para o *sistema* do segundo, sendo a *norma* o nível intermediário e, no campo da realização linguística, situam-se o *uso* hjelmsleviano e a *fala* coseriana.

Nesse sentido, norma aqui não é entendida como prescritiva, mas como eletiva, isto é, dentre tudo o que é possibilitado pelo sistema, uma comunidade escolhe o que lhe é mais adequado, sem, contudo, desviar-se do sistema. Isso explica, por exemplo, porque o português brasileiro é, por vezes, tão distinto de sua variante europeia, enquanto realização linguística, e mesmo assim seus sistemas são fundamentalmente os mesmos.

E essas particularidades apresentam-se especialmente no léxico da língua à medida que, dentre todo o repertório vocabular de uma língua, certos itens lexicais são eleitos para maior uso que outros, bem como alguns caem em desuso ou mesmo voltam à baila em outros momentos, ampliando sempre o repertório lexical, haja vista que “só existe uma possibilidade para um sistema lexical se cristalizar: a morte da língua. [...] Se a língua, porém, continuar a existir como meio

de comunicação oral (e também escrito), seu léxico se ampliará sempre” (BIDERMAN, 1978, p. 158).

Contudo, essas escolhas não são apenas realizadas individualmente pelos falantes, mas são norteadas pelo grupo em que esses usuários se inserem. A norma pode ser aquela ditada por grupos que gozem de maior prestígio na comunidade, excluindo, por exemplo, normas lexicais de outros falantes, por vezes menos prestigiados na sociedade, e a norma lexical coletiva, aquela aceita por um grande grupo, torna-se um padrão de comunicação, pelo qual as pessoas desse grupo conseguem interagir verbalmente. Porém, também os usuários, dentro dessa norma padrão, que é coletiva, fazem suas escolhas individuais, o que indica que um mesmo falante não faz uso de apenas um repertório lexical em todas as suas realizações, nem dois falantes de uma mesma comunidade partilham do mesmo vocabulário *ipsis litteris*. Isso porque, de acordo com as situações de comunicação, o usuário utiliza determinado léxico, para se adequar à comunidade em que está naquele momento ou mesmo para ser compreendido.

Tendo em vista o exposto, observa-se que há duas tensões que norteiam a norma: a ditada pelos usuários em nome de sua liberdade de escolha, e a gerada pela coletividade, em prol da natureza comunicativa, com vistas, pois, a uma relativa uniformidade.

Ainda com relação à norma, tendo em vista que o léxico nomeia os elementos dos ambientes físico e social de uma comunidade, o repertório vocabular registra particularidades de um momento social, histórico, geográfico e econômico de um dado grupo, o que pode aproximar a Lexicologia da Sociologia, e as palavras podem então ser entendidas como *testemunhas* de um momento sócio-histórico, como a tese defendida por Matoré (1953 *apud* OLIVEIRA, 1999, p. 53):

[...] cada um dos períodos históricos é caracterizado por palavras-testemunhas, ou seja, neologismos cujo nascimento, em um dado período e numa dada comunidade linguística, é determinado por situações de ordem social, econômica, totalmente novas.

Ainda nesse sentido, de acordo com Isquierdo (2003, p. 165), “o repertório lexical de uma comunidade linguística renova-se, transforma-se à medida que atuarem sobre ele fatores históricos, geográficos, culturais”. Então, além do momento histórico-cultural, também questões geoeconômicas influenciam a norma lexical de uma comunidade. Assim, “o léxico da língua é que mais nitidamente reflete o ambiente físico e social dos falantes” (SAPIR, 1969, p. 45).

E se o homem interage com seu ambiente, físico e social, os fatores extralinguísticos são expressos pelo léxico e também podem ajudar a desenhar a norma em uso pelos falantes, uma vez que precisam ser nomeados. Isso porque “é no âmbito de um grupo social que a norma se instaura, é disseminada ou fica confinada a determinados espaços geográficos, dependendo das condições socioculturais desse grupo” (ISQUERDO, 2006, p. 22).

Assim, determinados elementos culturais particulares de uma dada região geográfica só serão conhecidos e nomeados pela população que os conhecem, haja vista serem parte da realidade. Isso explica, por exemplo, como um mesmo animal pode receber designações diversas em áreas geográficas diferentes, nomes esses que dizem respeito a essa ou a aquela característica do animal, mais importante para o grupo social, ou ainda se pode entender o nome ainda desconhecido de um elemento cultural mencionado por um falante que não partilha da mesma norma linguística, advindo de outros espaços geográficos.

Destarte, a norma linguística varia em função de aspectos históricos, como já pontuado, e também em função de espaços determinados. No Brasil, por exemplo, várias são as normas em uso pelos falantes pertencentes a diversas

regiões político-administrativas. É sabido que o falar nordestino em muito difere do sulista, apesar de ambas as regiões terem sua norma linguística a partir do mesmo sistema. Levando-se em consideração o exposto, entende-se então que

[...] o espaço geográfico evidencia a particularidade de cada terra, exibindo a variedade que a língua assume de uma região para outra, como forma de responder à diversidade cultural, à natureza da formação demográfica da área, à própria base linguística preexistentes e à interferência de outras línguas que se tenham feito presentes naquele espaço no curso da história (CARDOSO, 2010, p. 15).

Portanto, tendo em vista o momento histórico que motivou o surgimento das *palavras testemunhas*, segundo Matoré (1953 *apud* OLIVEIRA, 1999), e também a influência de aspectos geográficos na eleição do léxico em uso aqui lembrados por Isquerdo (2006) e por Cardoso (2010), as exigências do ambiente físico e social serão molas propulsoras para a renovação lexical e determinantes para a formação da norma lexical em uso num dado momento por certo grupo de usuários.

Então, por vezes, necessidades advindas do momento sócio-histórico e determinadas pelos espaços físicos e sociais relativos a uma comunidade linguística exigem adequações, no nível lexical, que então se transmuta numa norma mais próxima e até adequada às situações comunicativas daquele momento, ocasionando, assim, mudanças linguísticas.

Dessa maneira, “as mudanças na língua seriam, em sua maioria, uma resposta às novas exigências de expressão, decorrentes das alterações no universo cultural da comunidade linguística, das descobertas científicas e tecnológicas etc.” (LUCCHESI, 1998, p. 95). Tais mudanças podem desencadear a já mencionada renovação lexical, tanto pela criação de novos itens lexicais quanto pelos novos sentidos atribuídos a palavras já existentes.

E se as línguas não são imunes às trocas sociais decorrentes dos contatos entre povos, e se a norma deve se adequar à realidade social e histórica da comunidade que a utiliza, essas relações refletem-se claramente no léxico por meio dos empréstimos linguísticos:

[...] há *empréstimo lingüístico* quando um falar A usa e acaba por integrar uma unidade ou um traço lingüístico que existia precedentemente num falar B e que A não possuía; a unidade ou o traço emprestado são, por sua vez, chamados de *empréstimos* (DUBOIS et al, 2006).

E esses empréstimos configuram-se como um aspecto que enriquece e caracteriza uma dada norma lexical, dando-lhe cor própria, decorrente da interação entre indivíduo, léxico e sociedade. Conforme argumenta Biderman (1978, p.139),

[...] qualquer sistema léxico é a somatória de toda a experiência acumulada de uma sociedade e do acervo da sua cultura através das idades. Os membros dessa sociedade funcionam como sujeitos-agentes, no processo de perpetuação e reelaboração contínua do Léxico da sua língua.

Também Mattoso Câmara Jr (1998, p. 104-105) define esses itens lexicais, de modo que sua definição os aproxima sobremaneira da realidade linguística brasileira, tendo em vista a abordagem sociocultural destacada no excerto a seguir:

[...] Empréstimo é a ação de traços diversos dos do sistema tradicional. O condicionamento social para os empréstimos é o contacto (sic) entre povos de línguas diferentes, o qual pode ser por coincidência ou posição geográfica, ou, à distância, por intercâmbio cultural em sentido *lato*. A coincidência ou posição geográfica determina os empréstimos íntimos e a língua a que é feito o empréstimo constitui um substrato, um superstrato ou um adstrato. Os empréstimos à distância são culturais.

No caso da relação entre as línguas indígenas e o português no Brasil, as primeiras são consideradas adstratos, tendo em vista que essas línguas eram faladas no mesmo momento histórico que o português, influenciando-o.

Os empréstimos linguísticos podem ser externos ou internos (CORREIA, 2008). O primeiro grupo engloba aqueles termos que foram trazidos de outras línguas; já o segundo diz respeito aos empréstimos dentro de uma mesma língua, como, por exemplo, quando a língua geral toma de empréstimo um termo específico de uma linguagem técnica.

Segundo Correia (2008), os empréstimos podem também ser lexicais ou semânticos. Lexicais se se tratarem de palavras e semânticos se forem novas acepções a palavras já utilizadas, sendo-lhes então emprestados alguns traços semânticos. Para a autora, “os empréstimos são, pois, uma consequência de situações de contato entre línguas diferentes, ou entre variedades e registros diferentes de uma mesma língua” (CORREIA, 2008, p. 41).

Já indigenismo, segundo Houaiss (2001), na rubrica Linguística, é definido como “palavra, construção ou locução de uma língua indígena tomada de empréstimo por outra língua”.

Então, a variante brasileira do português foi e continua sendo enriquecida e reelaborada pela presença de empréstimos linguísticos desde sua constituição, primeiro pela presença dos indigenismos, depois dos africanismos e, até os dias atuais, pelas diversas línguas dos povos que aqui aportaram em busca de uma nova história de vida, além dos contatos fronteiriços, entre outros.

2.2. APRESENTAÇÃO DO APORTE TEÓRICO

Este estudo teve como aporte teórico a Lexicologia, a Dialetoлогия, a Geolinguística, a Semântica e a Etnolinguística, áreas da Linguística focalizadas neste subtópico.

A Dialetoologia é o “ramo dos estudos da ciência da linguagem que se ocupa da variação e da diversidade de usos” (CARDOSO, 2008, p.16). Essa variação, bem como a diversidade de usos mencionada por Cardoso (2008) são motivadas por vários fatores, muitos deles extralinguísticos, como o espaço social e geográfico em que a língua se insere. O espaço geográfico

[...] evidencia a particularidade de cada terra, exibindo a variedade que a língua assume de uma região para outra, como forma de responder à diversidade cultural, à natureza da formação demográfica da área, à própria base linguística preexistente e à interferência de outras línguas que se tenham feito presentes naquele espaço no curso da história (CARDOSO, 2010, p.15).

No Brasil, a Dialetoologia teve como marco inicial o ano de 1826, quando Domingos Borges de Barros, o Visconde de Pedra Branca, a pedido de Adrien Balbi, participou do *Atlas Ethnographique du Globe*, contribuindo com um estudo de natureza dialetológica sobre o português do Brasil. Para Antenor Nascentes, segundo Mota e Cardoso (2006), essa seria a primeira fase da Dialetoologia brasileira, que se encerrou na década de 1920, quando se iniciou segunda fase, marcada pela publicação de *O dialeto caipira*, de Amadeu Amaral, que congregaria as pesquisas dialetológicas até os idos de 1950, quando Nascentes propôs a divisão.

Essa divisão é em parte aceita por pesquisadores da Dialetoologia em todo o País. Em parte porque novas pesquisas têm comprovado a existência de uma terceira e de uma quarta fases.

Ferreira e Cardoso (1994, p. 37-51), por exemplo, propuseram uma terceira fase, que se inicia a partir da publicação do Decreto Lei de número 30.643, de 20 de março de 1952. Então, a “mentalidade dialetológica” foi sendo sedimentada no Brasil, e tiveram início os trabalhos de natureza geolinguística, tanto os voltados

para a elaboração de atlas regionais como estudos de menores dimensões, como os atlas de pequenas regiões, dentre outros trabalhos.

A quarta fase, proposta por Mota e Cardoso (2006), reuniria os trabalhos de Dialetoлогия produzidos a partir de 1996, data do início do Projeto Atlas Linguístico do Brasil¹⁹, que agregam o viés da Geolinguística Pluridimensional, à medida que estende a motivação de suas pesquisas, da puramente diatópica, para considerar também outras variáveis.

Há ainda que se considerar a proximidade entre Dialetoлогия e Sociolinguística, que se dedica “*a estudiar sistemas en su contexto social*”²⁰ (HUDSON, 1994 *apud* RAMIREZ, 2009, p. 41). Essas outras variáveis demonstram a proximidade entre esses dois ramos da Linguística, uma vez que se considera que “*la variación entre un número determinado de elementos se correlaciona a variables sociales de los hablantes (estatus social, sexo, edad, profesión, etnia etc.)*”²¹ (RAMIREZ, 2009, p. 47).

Além disso, a Dialetoлогия também se vale das contribuições de outras ciências, especialmente as sociais, como a história, a antropologia, a etnologia, por exemplo, para analisar as informações catalogadas, ratificando que “a dimensão social da língua é tão importante quanto a espacial” (ISQUERDO, 2005, p. 335).

Para isso, a Dialetoлогия vale-se também do método fornecido pela Geolinguística ou Geografia Linguística, que “como la Dialectologia, estudia la variación de la lengua, pero la estudia en el espacio, e la refleja sobre mapas”²² (MOUTON, 2009, p. 63). Porém, a Geolinguística contemporânea considera também

¹⁹ Cf. Capítulo III deste trabalho.

²⁰ “[...] a estudar sistemas em seu contexto social” (TN).

²¹ “A variação entre um número determinados de elementos se correlaciona a variáveis sociais dos falantes (status social, sexo, idade, profissão, etnia etc)” (TN).

²² “[...] como a Dialetoлогия, estuda a variação da língua, mas a estuda no espaço e se reflete sobre mapas” (TN).

outros fatores para a descrição da língua, como atestam Mota e Cardoso (2006, p. 22):

[...] Na geolinguística pluridimensional contemporânea, soma-se ao parâmetro diatópico, prioritário em trabalhos dessa natureza, o interesse por outros tipos de variação, como a diagenérica [variação entre os gêneros/sexos masculino e feminino], a diastrática [variantes sociais], a diageracional [considerando-se a idade dos falantes], entre outras.

O uso das terminologias Geografia Linguística e Geolinguística é esclarecido por Mouton (2009, p. 64), para quem a

[...] Geografía Lingüística traduce la denominación francesa *géographie linguistiques*, que fue la primera em utilizarse, y se tradujo también a las demais lenguas românicas. [...] desde hace años, também geolinguística se emplea como sinónimo, sobre todo entre los romanistas²³.

A Geografia Linguística é, pois, fundamental para o mapeamento dos dados coletados por pesquisas dialetológicas, uma vez que registra as informações, por meio de cartas/mapas linguísticos que fornecem recursos práticos para a análise do fenômeno linguístico mapeado:

[...] o registro em mapas especiais de um número relativamente elevado de formas linguísticas (fônicas, lexicais ou gramaticais) comprovadas mediante pesquisa direta e unitária numa rede de pontos de um determinado território, ou que, pelo menos tem em conta a distribuição das formas no espaço geográfico correspondente à língua, às línguas, aos dialetos ou aos falares estudados (COSERIU, 1982, p.79).

Além da Dialetologia e da Geolinguística, este trabalho tem suas bases também na Lexicologia, haja vista ser um estudo de natureza lexical. A Lexicologia, ramo das ciências do léxico que estuda a origem e as transformações fonéticas, morfológicas, sintáticas e semânticas das palavras, comparando-as com outras da mesma língua ou de línguas diversas, numa perspectiva sincrônica ou diacrônica,

²³ “Geografia Linguística traduz a denominação francesa *géographie linguistiques*, que foi a primeira a se utilizar, e se traduziu também para as línguas românicas. [...] Durante anos, também geolinguística se emprega como sinônimo, sobretudo entre os romanistas” (tradução nossa).

sustenta teoricamente este trabalho uma vez que se associam seus conceitos aos da Dialetologia, aos da Geolinguística e também aos da Semântica, à medida que o estudo aqui proposto concebe o léxico como produto social.

Intenta-se, pois, estudar até que ponto o contato entre as sociedades reflete-se no acervo vocabular dos habitantes das capitais brasileiras, por meio da investigação dos empréstimos linguísticos – indigenismos – identificados no *corpus* tomado como fonte.

Assim, foca-se a dimensão etnolinguística do estudo, que busca investigar relações entre as culturas que participam e participaram da formação da identidade sociocultural brasileira e seus desdobramentos na variante brasileira do português.

Toma-se aqui a vertente teórica de que a língua reflete os aspectos culturais das sociedades que a utilizam, o relativismo linguístico proposto por Sapir, uma vez que “toda língua tem uma sede (uma comunidade). O povo que a fala pertence a uma raça ou a certo número de raças, isto é, a um grupo de homens que se destaca de outros grupos por caracteres físicos” (SAPIR, 1969, p. 205). Nessa perspectiva, a língua é concebida como intrinsecamente relacionada à cultura e à sociedade a que pertence, revelando aspectos socioculturais de uma determinada época e mesmo tornando-se ela produto cultural, dado o fato de estar profundamente relacionada ao homem em relação a seu ambiente social.

Reproduzindo-se a posição do mesmo linguista, pode-se entender assim “que o léxico reflita em alto grau a complexidade da cultura é praticamente um fato de evidência imediata [...] destina-se em qualquer época a funcionar como um conjunto de símbolos, referentes ao quadro cultural do grupo” (SAPIR, 1969, p. 51).

O campo de estudos da Etnolinguística foi definido por Sapir (1971), em seu livro *A linguagem: Introdução ao estudo da fala*, em que dedica um capítulo à discussão da relação entre língua, raça e cultura. Já Coseriu (1978) ocupou-se da definição das tarefas da Etnolinguística, por entender que seria necessário repartir as pesquisas em várias disciplinas, dada a amplitude da sua área de trabalho. Para esse linguista, a Etnografia Linguística seria a responsável por analisar a relação entre a língua e a cultura, e trataria dos saberes acerca das realidades expressos pela linguagem.

No Brasil, os estudos etnolinguísticos foram inicialmente empreendidos, especialmente por missionários religiosos, além de europeus e depois por brasileiros, muitos deles sem formação linguística. A comissão Rondon, por exemplo, foi responsável pela documentação de diversas línguas indígenas do Brasil. O Marechal Cândido Mariano Rondon estava a serviço do governo brasileiro para instalar as linhas telegráficas que facilitariam a comunicação em terras brasileiras, desde 1892, razão pela qual acabou por entrar em contato com etnias isoladas e participando de suas vidas por algum tempo, o que resultou na documentação de línguas indígenas, que dão base a estudos linguísticos até a atualidade.

Todavia, por também serem realizadas por não linguistas, e especialmente pela visão etnocentrista que norteou todas as pesquisas etnolinguísticas da época, as pesquisas dos séculos XIX e início do século XX pecavam, por vezes, pela falta de profundidade e de critérios linguístico-comparativos, o que possibilitaria uma melhor análise a partir dos registros realizados. Em resumo sobre essa época, pode-se entender que

[...] O tipo de pesquisa que predominou – o de levantar o vocabulário da língua debaixo de uma série padrão de itens de interesse

etnológico – teve pelo menos a vantagem de permitir um exame comparativo amplo, porque de um a outro pesquisador vamos encontrar sempre os mesmos itens semânticos. Foi uma evidente melhoria em relação a [Carl Friedrich Philipp von] Martius, que ficava obtendo a esmo os vocábulos, conforme as possibilidades de coleta²⁴. Desse trabalho, profundo e sério, apesar de suas deficiências, resulta, pois, uma base teórica para a compreensão das línguas indígenas e a tarefa de as classificar (MATTOSO CÂMARA JÚNIOR, 1977, p. 127).

Atualmente, os estudos etnolinguísticos continuam enriquecendo o conhecimento acerca das línguas indígenas brasileiras, mas também se ocupa do estudo de outras línguas, no que tange à sua relação com a cultura. Um reconhecido teórico brasileiro nessa área é Aryon Dall’Igna Rodrigues, pesquisador responsável pela criação e direção do Laboratório de Línguas Indígenas (LaLI) da Universidade de Brasília (UnB). Segundo ele,

[...] O objetivo maior do Laboratório de Línguas Indígenas do Instituto de Letras da Universidade de Brasília é o estabelecimento de um espaço institucional para promover a documentação, análise, descrição, comparação não só das línguas, mas também das situações em que se encontram estas (RODRIGUES, 2012a. Disponível em: <http://www.comciencia.br/reportagens/linguagem/ling13.htm>).

Além disso, Rodrigues é responsável pela publicação de diversas obras, muitas delas que serviram de base para este trabalho, e também orientação de vários estudos de pós-graduação na área, que focam a presença da herança indígena no vocabulário do português do Brasil.

Há também outro núcleo de referência em estudos das línguas indígenas no Brasil, sediado no Museu Nacional e que tem, dentre outros pesquisadores, Yonne Leite. Fundado em “1958, com a participação dos professores Luís de Castro Farias, Darcy Ribeiro e Joaquim Mattoso Câmara Jr., [esse centro] tem por objetivo a documentação e a análise das línguas indígenas brasileiras e das variantes do

²⁴ Von Martius foi o responsável pela estruturação do estudo do tronco Jê e elucidou o *melting-pot* do grupo Tapuia, dentre várias outras contribuições aos trabalhos etnolinguísticos no Brasil (MATTOSO CÂMARA JR, 1977).

português do Brasil” (MUSEU NACIONAL. Disponível em: <http://www.museunacional.ufrj.br/linguistica/historico.htm>).

Este estudo também espera colaborar com as pesquisas etnolinguísticas brasileiras, uma vez que se propôs a estudar os indigenismos presentes nas áreas semânticas já mencionadas.

E, concluindo este subtópico que trata das áreas de Linguística que dão base a este trabalho, tem-se a Semântica, que também fornece aporte para as análises realizadas, haja vista que todas as unidades lexicais a serem analisadas têm considerados seus traços semânticos, sua significação, para que se possam ainda verificar, além das variações diagenéricas, diageracionais, diastráticas e diafásicas na realização do léxico, as inovações, alternâncias e especificidades semânticas registradas na nomeação dos conceitos expressos pelas perguntas do Questionário Semântico-Lexical do Projeto ALiB, selecionadas para o estudo.

Historicamente, a Semântica linguística nasceu sob a influência das ideias darwinistas, nos idos de 1880, tendo como principal mentor Michel Bréal, e intentava verificar o princípio metafórico evolucionista das línguas por meio do estudo das variações e mudanças ocorridas com as palavras e, conseqüentemente, com as frases e enunciados. Neste trabalho, utiliza-se a Semântica Lexical, que estuda “o sentido das palavras” (GUIRAUD, 1995, p. 05).

O termo palavra é bastante controverso no campo dos estudos lexicais. Todavia, apropriamo-nos aqui da definição de Irène Tamba, para quem, “numa primeira aproximação, as palavras são um conjunto de formas fônicas e gráficas autônomas, estáveis, que constituem o léxico de uma língua” (TAMBA, 2006, p. 70). Sendo então as palavras, no plano semântico, “os nomes, variáveis segundo as línguas, dados aos seres, coisas ou acontecimentos e que servem para identificá-los

e diferenciá-los” (TAMBA, 2006, p. 71), este estudo se pauta na Semântica para analisar o *corpus* pesquisado – designações para elementos nacionais auferidas a partir do banco de dados do Projeto ALiB.

2.3 O PORTUGUÊS DO BRASIL: ACHEGAS A UM ESTUDO ETNOLINGUÍSTICO

A história do português do Brasil pode ser entendida como a história de outras línguas de cultura, uma vez que “a história de uma língua realmente se esclarece pela história social e política do povo que usa essa língua” (MATTOS E SILVA, 2004, p. 91). No Brasil, os contatos interétnicos são marcantes para a reelaboração lexical do português, processo pelo qual influências linguísticas e extralinguísticas marcam a criação de novas unidades léxicas para nomear novos elementos, colocam em desuso certos itens lexicais, bem como associam novos sentidos aos já existentes.

O contato do colonizador português com os autóctones que aqui viviam, com os negros africanos trazidos como escravos e, posteriormente, com os demais imigrantes que para esta terra se deslocaram em busca de melhores condições de vida resultou em trocas culturais, refletidas nos hábitos, nas crenças e nas tradições que esses povos partilharam. E essas relações interétnicas estão presentes também na língua portuguesa do Brasil, especialmente pela presença considerável de unidades lexicais de base indígena e africana, que singularizam a diversidade da variante brasileira do português comparada ao português europeu.

Os contatos interétnicos no Brasil tiveram início no primeiro decanato do século XVI, com a chegada dos portugueses. Encantados com uma realidade totalmente diferente da europeia, já que o Novo Mundo abrigava matas fechadas,

grandes árvores e animais nunca vistos, os europeus encontraram na costa brasileira a maior etnia indígena habitante dessas terras: os índios tupi.

As relações entre o branco europeu e o autóctone, além do contato com os africanos, ajudaram a desenhar a formação social e cultural do povo brasileiro, à medida que favoreceram intercâmbios culturais e linguísticos entre essas etnias.

Estreitando-se os laços dos europeus com os autóctones, também os casamentos entre índias e portugueses passaram a ser benquistos na sociedade indígena, ocasionando a miscigenação tão característica da nação brasileira. Assim, as famílias, constituídas em sua maioria pela união do europeu com a índia, usavam como forma de comunicação a língua indígena, ficando o uso do português reservado para os documentos oficiais e seu ensino a cargo das escolas, segundo Vieira (1951 *apud* NAVARRO, 1998, p.174). Logo, o português era utilizado pela elite escolarizada da época colonial – em torno de 0,5% da população, segundo Mattos e Silva (2004), enquanto a Língua Geral (LG) era falada pelo restante da população, portugueses iletrados, brasileiros mestiços, índios e africanos.

Além disso, como já mencionada, considerando-se que o Brasil Colonial caracterizava-se como colônia de exploração de Portugal, era necessário que os portugueses aqui instalados relatassem à Coroa as riquezas do novo mundo, a nova realidade encontrada. E, para isso, por vezes, o vocabulário do europeu tornava-se insuficiente, carecendo de designativos para a nova realidade encontrada – animais, plantas e hábitos antes desconhecidos, o que ocasionou a designação de elementos tipicamente nacionais com nomes indígenas: “a diversidade étnica dos grupos populacionais brasileiros [...] acabou provocando o exercício de trocas ou empréstimos vocabulares, para nos atermos, aqui, apenas ao campo léxico” (DICK,

2000, p.305). Com a intensificação dos contatos e o crescimento da população mestiça,

[...] No transcorrer dos primeiros séculos de colonização, nos diferentes núcleos populacionais, surgia uma variante lingüística que permitia a comunicação entre índios, negros e mestiços que tinham de aprender a falar a língua da minoria branca por uma questão de sobrevivência (ISQUERDO, 2006, p. 12).

Essa variante lingüística passou a ser a língua falada em toda a Colônia, a Língua Geral (LG), em detrimento ao uso da língua transplantada. A LG recebeu essa denominação, segundo Rodrigues (2002, p. 100), apenas após o terceiro século da colonização, pois, ao longo do século XVII, era denominada de “língua brasílica” e depois de “tupinambá” (século XVIII), apesar de já ser conhecida por tal nome desde o século XVII. Tratava-se de uma língua de base indígena (tupi) com certa presença do português europeu.

[...] Em termos históricos, a expressão língua geral refere-se ao processo lingüístico e étnico instaurado no Brasil pelo complexo catequético-colonizador, cujo emprego aponta para três concepções: a) em sentido genérico, diz respeito às línguas surgidas na América do Sul em consequência dos contatos entre agentes das frentes de colonização e os grupos indígenas; b) especificamente, designa as línguas, de base indígena, desenvolvidas e instituídas em São Paulo e na Amazônia, e faladas por uma população supraétnica; c) refere-se também à gramaticalização dessas línguas ditas gerais (BORGES, 2001, p. 211).

Dessa forma, e por ser também utilizada na comunicação entre os mestiços e indígenas, integrantes das bandeiras²⁵, a Língua Geral passou a ser disseminada por todo o território nacional, recém-descoberto, sendo frequente o uso de designativos tupis para a nomeação de diversos acidentes geográficos e humanos, por exemplo, mesmo em áreas onde nunca se registrou a efetiva

²⁵ As bandeiras foram expedições que contribuíram sobremaneira para a expansão territorial brasileira, entre os séculos XVI e XVIII, haja vista adentrarem em terras desconhecidas (mais ao centro do território) em busca de metais preciosos e escravos, para além do domínio português, por vezes. Cf. Capítulo I deste trabalho.

presença da etnia tupi, conforme atesta Rodrigues (2002, p.21), ao salientar que “é notável a quantidade de lugares com nomes de origem Tupinambá, quase sem alteração de pronúncia, muitos deles dados por luso-brasileiros dos séculos passados a localidades onde nunca viveram índios Tupinambá”.

Segundo Mattos e Silva (2004), essa Língua Geral, genérica e no singular, foi a versão jesuítica da língua brasilica e funcionava como “um veículo fundamental como instrumento de interação linguística entre os portugueses [...] e os indígenas predominantemente do tronco linguístico tupi, e, na Capitania da Bahia” (MATTOS E SILVA, 2004, p. 77).

É mister identificarmos a LG haja vista que, segundo Rodrigues (2002), teríamos, como já mencionado, de fato, três línguas gerais: a paulista, base para o dialeto caipira e profundamente expandida em território nacional, por ser a língua dos bandeirantes paulistas; a amazônica, muito mais documentada e que ainda hoje é utilizada na região Amazônica e o Guaraní Criollo. Houaiss (1985, p. 49-50) cita ainda uma língua geral de base cariri, que se estenderia pelo Nordeste brasileiro e interior, à época colonial, todavia, pouco documentada.

Essa Língua Geral Paulista (RODRIGUES, 2002) dominou o cenário linguístico nacional por aproximadamente três séculos, e seu predomínio só foi contido após decreto da Metrópole Portuguesa, que via no uso da língua local uma ameaça para o controle da Colônia, na segunda metade do século XVIII. Em 1757, pela publicação do documento conhecido como Diretório dos Índios, o Marquês de Pombal, então ministro português, decretou que não mais fosse utilizada a língua híbrida em terras nacionais. Em consequência disso,

[...] O Diretório intervém, deste modo, no processo de construção da identidade linguística brasileira, já que visa impedir a brasilidade de enunciar-se de dentro da formação discursiva que lhe é própria. Esse Diretório vigorou durante aproximadamente 40 anos, sendo abolido

apenas em 12 de maio de 1798. De qualquer forma, as raízes desse português-brasileiro continuaram resistindo e produzindo efeitos no dizível possível, embora a língua geral e as demais línguas indígenas tivessem sofrido um processo sistemático de eliminação ao longo dos séculos seguintes (MARIANI, 2001, p. 111).

Além disso, a língua de base indígena passava então a ser menosprezada pela sociedade da época, em relação à língua da metrópole. Segundo ideias disseminadas a partir desse diretório, “a língua indígena era pobre, simples, rudimentar, sem passado literário, era instrumento de um povo rude, de cultura primitiva. Ao invés, o português era rico, complexo, maleável, possuidor de uma tradição literária” (MELO, 1981, p. 51).

Mesmo assim, e tendo decorridos mais de 500 anos desde o descobrimento, a presença tupi mantém-se no léxico da variante brasileira do português, por meio da representativa quantidade de designativos dessa origem, especialmente em áreas semânticas que nomeiam referentes tipicamente nacionais, como a fauna, a flora, a alimentação e as atividades ligadas ao campo.

Mattos e Silva (2004), inclusive, acrescenta que essa influência indígena, além da africana, deu bases para uma nova língua portuguesa, diversa da variante europeia transplantada para o Brasil, o português geral brasileiro, que modernamente gerou a variante popular da língua portuguesa no Brasil. Segundo a autora,

[...] Muito provavelmente a língua de comunicação generalizada nesse momento (séculos XVII e XVIII) integrador da sociedade colonial seria o português geral brasileiro, não se negando que em situações específicas pudessem aí ser ouvidas tanto línguas gerais indígenas, línguas indígenas, línguas africanas mais gerais entre os escravos [...] e, certamente, o português europeu, sobretudo porque um contingente significativo de portugueses se deslocou para a área mineradora em busca de ouro (MATTOS E SILVA, 2004, p. 86).

Assim, além da tupi, devem-se ressaltar várias outras etnias que participaram da formação da nação e exerceram influências na configuração da

norma linguística brasileira, em maior ou menor proporção, de acordo com o grau de contato entre os povos, além de sua localização geográfica. Os negros, como já mencionado, além de outros indígenas, como os Guaicuru, os Guató e os Terena, no Centro-Oeste, os Ianomâmi, os Parintins e os Xavante no Norte e os Potiguara, no Nordeste, para se citarem apenas alguns exemplos.

Contudo, a contribuição das demais línguas indígenas no léxico do português do Brasil se processou de forma bem mais tímida que a herança tupi. De acordo com Barbosa ([s.d.] *apud* HOUAISS, 1985, p. 153), “a contribuição tapuia²⁶ será menor que a chinesa, persa, etc., [...] visto como aquelas línguas tiveram contacto muito passageiro e circunscrito a pequenas regiões”.

Na verdade, toda a contribuição indígena ao léxico do português do Brasil traduziu-se na presença de indigenismos, vocábulos de base indígena incorporadas à língua nacional. Segundo Melo (1981), esses indigenismos caracterizam, ao lado dos africanismos e das demais influências étnicas de cunho lexical, a variante brasileira do português. E essas unidades léxicas que caracterizam o português do Brasil, segundo o mesmo autor, são os chamados brasileirismos. Assim, para Melo (1981), os brasileirismos são o que difere e particulariza a língua falada no Brasil em relação àquela do continente europeu, e seriam divididos em seis categorias, de acordo com sua origem e função:

1. Tupinismos
2. Africanismos
3. Vozes ameríndias e hispano-americanas
4. Formações e derivações brasileiras
5. Brasileirismos quanto à significação
6. Arcaísmos (MELO, 1981, p. 149).

²⁶ A designação tapuia foi cunhada a partir da primeira classificação das línguas indígenas, distribuídas entre Tupi e Tapúya, pelos primeiros missionários e colonizadores (MELATTI, 1980, p. 33). Essa nomenclatura tem sido refutada, uma vez que sugeriria uma homogeneidade entre as línguas não tupi, fato que já foi comprovadamente derrubado pelos estudos linguísticos de temática indígena.

Interessam-nos especialmente, para este trabalho, os itens 1 e 3, que dizem respeito aos indigenismos. O autor em questão separa em dois itens as influências indígenas tendo em vista a representatividade desses segmentos na língua portuguesa brasileira. Isso porque os tupinismos, unidades léxicas de base linguística tupi, estão presentes em quantidade consideravelmente maior na língua que as demais palavras indígenas, como, inclusive, já fora observado por Barbosa ([s.d.] *apud* HOUAISS, 1985, p. 153).

Ainda sobre os indigenismos, Oliveira (1999), considerando a proposta de Melo (1981), retoma as posições de D’Albuquerque, que assim sistematiza os brasileirismos de base indígena:

1. Palavras originárias de línguas indígenas e incorporadas ao nosso léxico. [...]
5. Vocábulo de outras línguas que aqui se deturpam (D’ALBUQUERQUE, [s.d.] *apud* OLIVEIRA, 1999, p. 64).

Nessa separação, Oliveira (1999, p. 65) considera que, de forma implícita, as línguas indígenas a serem consideradas no item 1 proposto por D’Albuquerque seriam “apenas as brasileiras”, inserindo-se no quinto item as palavras originadas das demais línguas indígenas.

Segundo Rodrigues (2002, p.18), “falam-se no Brasil, hoje em dia, umas 170 línguas indígenas”, quantidade consideravelmente inferior ao montante de aproximadamente mil línguas, à época do início da colonização. Essa diminuição no número de línguas indígenas em uso no Brasil pode ser explicada pelo próprio desaparecimento físico dos falantes

[...] em decorrência de epidemias, extermínio direto, escravização, redução de territórios, destruição das condições de sobrevivência e aculturação forçada, entre outros fatores que sempre acompanharam as frentes de expansão desde o período colonial até nossos dias (SEKI. *Línguas indígenas do Brasil no limiar do século XXI*. Disponível em <http://www.unimep.br/phpg/editora/revistaspdf/imp27art11.pdf>).

Em determinadas regiões, a população indígena e, conseqüentemente, suas línguas, foram praticamente dizimadas, enquanto em outras se mantiveram com certa representatividade, situação motivada por questões econômicas, como argumenta Rodrigues (2002, p.19), ao esclarecer que “naturalmente, o maior número de línguas indígenas desapareceu nas áreas que foram colonizadas há mais tempo e mais intensamente, constituídas pela região Sudeste e pela maior parte das regiões Nordeste e Sul do Brasil”.

As línguas indígenas brasileiras, em sua maioria, são provenientes de dois troncos linguísticos, Tupi e Macro-Jê, que mantêm sua permanência na variante brasileira do português especialmente em áreas onde a colonização foi tardia ou não intensa, como nos casos das regiões Norte e Centro-Oeste do Brasil, áreas que, segundo Prado Jr (1986²⁷, p.113), na década de 1940, eram “em geral as de vida econômica pouco ativa”.

2.4 O LEGADO INDÍGENA NO LÉXICO DO BRASIL: DA LÍNGUA GERAL AOS DIAS ATUAIS

Desde o início da colonização brasileira, o elemento indígena tem marcado presença na história social da nação, pelos seus hábitos, tradições e crenças transmitidos ao longo das gerações. Como produto social, a língua também reflete essa influência, uma vez que

[...] todos os componentes de uma língua – seu sistema de sons, seu sistema morfológico e sintático e seu vocabulário, assim como suas estratégias de construção do discurso – mudam no curso do tempo, em razão de reajustes internos desses sistemas e devido a mudanças na cultura e organização social do povo que a fala e a

²⁷ A primeira edição da obra *Formação do Brasil contemporâneo*, de Caio Prado Júnior, é de 1942. Para este trabalho consultamos a 19ª edição, publicada em 1986.

influências de outras línguas com que ela entra em contacto em determinadas circunstâncias (RODRIGUES. As línguas gerais sul americanas, disponível em http://vsites.unb.br/il/lali/publicacoes/publ_002.html).

O legado indígena permanece, assim, na cultura e na norma linguística brasileira, mesmo passados mais de quinhentos anos desde o descobrimento, especialmente em regiões em que o progresso econômico e industrial chegou mais lentamente, como no Norte e no Centro-Oeste do Brasil. Nessas duas regiões político-administrativas, inclusive, estão localizados os dois Estados que abrigam as maiores populações indígenas do Brasil: Amazonas e Mato Grosso do Sul (IBGE, 2012).

Também no Norte e no Centro-Oeste do Brasil ainda é possível encontrar grande número de falantes de línguas indígenas, registrando-se também a coexistência oficial de línguas indígenas com o português do Brasil, como em São Gabriel da Cachoeira/AM (*Nheengatu*, *Tukano* e *Baniwa*) e em Tacuru/MS (*Guarani*). Rodrigues (2002, p.33), nesse contexto, relata que no Norte e no Centro-Oeste ainda se registram diversas línguas do tronco Tupi em atividade, sendo,

[...] ao todo, 21 línguas vivas da família Tupi-Guarani, que identificamos em território brasileiro [...] faladas por cerca de 33.000 pessoas. Contam com maior número de falantes o Kaiwá, em Mato Grosso do Sul, e o Tenetehára (Guajajára e Tembé), no Maranhão e Pará, com cerca de 7.000 cada um. O Kaiwá e o Tenetehára são as línguas tupi-guarani mais populosas do Brasil.

Já em relação ao Macro-Jê, Rodrigues (2002, p. 47) argumenta que

[...] o constituinte maior do tronco Macro-Jê é a família linguística Jê, que compreende línguas faladas sobretudo nas regiões de campos cerrados que se estendem ao sul do Maranhão e do Pará, em direção ao sul, pelos estados de Goiás e Mato Grosso, até os campos meridionais dos estados de São Paulo, Paraná, Santa Catarina e Rio Grande do Sul.

A par das línguas derivadas desses dois troncos linguísticos, registram-se também em localidades dessas regiões outras famílias linguísticas, como a Karib e a Aruák, além de diversas outras línguas não associadas a qualquer família linguística.

Frente ao exposto, vê-se o seguinte quadro em termos das línguas indígenas vivas no Brasil:

[...] línguas da família Tupí-Guaraní no Brasil, que é a mais distribuída sobre nosso território, com línguas no Amapá e norte do Pará e com outras no Paraná, Santa Catarina e Rio Grande do Sul, com umas no litoral atlântico e outras em Rondônia, assim como nos principais afluentes meridionais do rio Amazonas, no Madeira, no Tapajós, no Xingu e também no Tocantins e Araguaia. Outras grandes famílias são a Jê, que tem línguas distribuídas desde o Maranhão até o Rio Grande do Sul, a Aruak no oeste e no leste da Amazônia, em Mato Grosso e em Mato Grosso do Sul, e a Karib ao norte do rio Amazonas, nos estados do Amazonas, Roraima, Pará e Amapá, mas com algumas línguas ao sul daquele rio, ao longo de seu afluente Xingu, nos estados do Pará e Mato Grosso (RODRIGUES, 2012a).

E essa presença influencia as escolhas lexicais dos brasileiros, nas diversas regiões, com a manutenção de indigenismos na designação de elementos da cultura nacional. Porém, além do fator histórico e de contatos étnicos, devem-se considerar ainda os elementos geográficos no desenhar da norma lexical brasileira, como as fronteiras secas com países da América do Sul com forte influência indígena, como Colômbia, Peru, Argentina, Bolívia e Paraguai, por exemplo, fator que também pode contribuir para a manutenção do substrato indígena na fala dos brasileiros. Esses países valorizam a importância indígena, traduzida, por exemplo, pelas línguas co-oficiais de origem indígena – como o quíchua no Peru e na Bolívia, que comporta ainda o aimará, e o guarani no Paraguai. E, como as trocas culturais são intensas entre esses países e o Brasil, e a língua é também produto cultural, a comunicação entre seus habitantes comporta palavras “emprestadas” das línguas vizinhas.

Um exemplo que ilustra essa situação de contato é o comércio entre as cidades de Ponta Porã, em Mato Grosso do Sul, e Pedro Juan Caballero, no Paraguai, em que transitam as línguas espanhola, portuguesa e guarani, mesmo, por vezes, a língua guarani funcionando como uma proteção para seus falantes, excluindo aqueles que a desconhecem.

Em síntese, a presença de línguas indígenas no português brasileiro perpassa a era colonial e os séculos subsequentes, transitando ainda na atualidade em diversos pontos do Brasil, o que está sendo investigado no âmbito deste trabalho, que busca analisar até que ponto a proximidade/distância dessas áreas híbridas, linguisticamente, influencia a norma linguística dos habitantes de grandes centros urbanos, no caso, das 25 capitais brasileiras que integram a rede de pontos do Projeto ALiB, bem como verificar se os casos documentados fazem parte da norma padrão ou revelam peculiaridades regionais.

Sendo assim, selecionou-se um percurso metodológico para se chegar a esse objetivo, que será apresentado no capítulo subsequente, o Capítulo III, que versa sobre as obras pesquisadas e a forma com que foram analisadas as unidades léxicas coletadas a partir dos dados catalogados pelo Projeto ALiB, que também será contextualizado no capítulo a seguir. A ferramenta tecnológica que propiciou o tratamento das informações também está detalhada no Capítulo III, dentre outras informações de cunho metodológico.

CAPÍTULO III

METODOLOGIA DA PESQUISA

Os princípios teórico-metodológicos que nortearam este trabalho foram buscados especialmente na Dialetoлогия, na Geolinguística, na Lexicologia e na Semântica, e a dimensão do trabalho é de natureza etnolinguística. Também a Dialetoлогия e a Geolinguística Pluridimensional foram o norte para o projeto que forneceu os dados e também o caminho metodológico para a coleta dos dados utilizados neste estudo, o Projeto Atlas Linguístico do Brasil.

Como o Projeto ALiB, na coleta dos dados, controla as variáveis espaço geográfico, idade, sexo, escolaridade e situação social dos falantes, os dados são analisados segundo diferentes dimensões, como a diatópica e a léxico-semântica: de forma geral, objetivou-se analisar traços das línguas indígenas presentes no léxico do português do Brasil a partir de dados do Projeto Atlas Linguístico do Brasil, em 25 capitais brasileiras. E, como já mencionado na Introdução, intentou-se também identificar isoglossas léxicas a partir da distribuição diatópica dos itens lexicais estudados, distribuição essa que poderá também auxiliar o registro de marcas de uso presentes nas obras lexicográficas, e se procurou relacionar a presença/ausência das etnias indígenas nesses espaços.

Como já assinalado, o *corpus* deste estudo compõe-se de dados geolinguísticos coletados por pesquisadores vinculados ao Projeto Atlas Linguístico do Brasil, documentados com o auxílio do questionário linguístico elaborado para o Projeto em questão (2001), que visa a catalogar a dinamicidade da língua em uso

pelos brasileiros. Neste caso, são considerados os dados relativos aos indigenismos identificados na fala dos habitantes entrevistados.

Embora haja outros substratos no PB (Português do Brasil), decidiu-se pelo estudo dos indigenismos por esses serem provenientes de línguas de fato autóctones em território nacional, e cujos estudos já vêm demonstrando uma presença substancial no léxico especialmente de áreas semânticas que reúnem itens lexicais que nomeiam referentes, em sua maioria, mais específicos da realidade brasileira.

Essa constatação também iluminou a eleição das três áreas semânticas do Questionário Semântico-lexical (ALiB/2001) contempladas por este estudo: *atividades agropastoris, fauna e alimentação e cozinha*.

Como a base de dados é pertencente ao Projeto ALiB, essas áreas semânticas totalizam, no questionário semântico-lexical do referido projeto, um montante de 62 questões (25 relativas às *atividades agropastoris*; 25 relacionadas à *fauna* e 12 que dizem respeito à *alimentação e cozinha*), que geraram respostas que foram analisadas inicialmente quanto à base linguística.

Para isso, foram consultados dicionários de língua portuguesa e de línguas indígenas contemporâneos, dicionários históricos e, especialmente, dicionários etimológicos, para fins de seleção das unidades léxicas de base linguística indígena. Dentre os dicionários gerais de língua portuguesa, foram consultadas as versões eletrônicas do *Dicionário Eletrônico Houaiss da Língua Portuguesa* (HOUAISS, 2001) e do *Novo Dicionário Aurélio da Língua Portuguesa* (FERREIRA, 2004). E, entre as obras de cunho etimológico, figuram o *Dicionário etimológico da Língua Portuguesa*, de José Pedro Machado (1987), além do *Dicionário etimológico da língua português*, de Antônio Geraldo da Cunha (2010).

De cunho indígena, foram consultadas as obras de Sampaio (1928), *O tupi na geographia nacional*, e Cunha (1988), *Dicionário histórico das palavras portuguesas de origem tupi*, como base de pesquisa.

Outras obras foram também consultadas quando da necessidade de maiores esclarecimentos quando à origem, acepção ou mesmo registro das unidades léxicas aqui estudadas.

A partir desse primeiro escrutínio, foi constituído o *corpus* do estudo que reuniu 59 unidades léxicas de base indígena, que foram analisadas também em termos diatópico e léxico-semântico, com destaque para traços indígenas no vocabulário em uso pelos brasileiros habitantes das capitais, hipótese que norteou este estudo.

Dada a importância do Projeto ALiB para este estudo, por fornecer os dados analisados e o norte para a metodologia adotada, como também pelas contribuições desse projeto para os estudos linguísticos, e ainda seu caráter geolinguístico, o próximo subtópico foi reservado para a caracterização desse Projeto.

3.1. PROJETO ATLAS LINGUÍSTICO DO BRASIL: REVELANDO FALARES, DESCORTINANDO A LÍNGUA EM USO

A proposta de um atlas linguístico nacional do português do Brasil foi lançada oficialmente em 20 de março de 1952, quando da publicação do Decreto-Lei 30.643, em que o governo brasileiro determinou, como principal finalidade da Comissão de Filologia da Casa de Rui Barbosa, a elaboração de um Atlas linguístico do Brasil.

Todavia, essa proposta de um atlas nacional foi retomada apenas em 1996, em Salvador, Bahia, por ocasião do Seminário “Caminhos e perspectivas da geolinguística no Brasil”, quando tomou forma não só a proposta do texto legal quanto a ideia plantada por filólogos e linguistas como Serafim da Silva Neto, Antenor Nascentes, Celso Cunha, que defendiam a proposta de se iniciarem os estudos geolinguísticos no Brasil pelos atlas regionais, o que se concretizou e teve como precursor inicial o trabalho coordenado por Nelson Rossi, que deu origem ao *Atlas Prévio dos Falares Baianos* (1963). Outros pesquisadores deram continuidade à execução de atlas regionais, que contribuíram e contribuem ainda hoje com suas experiências, no que diz respeito à metodologia e especialmente ao viço e à paixão pela pesquisa dialetológica.

Esse intervalo de tempo, de quase cinquenta anos, entre o tema ter sido matéria de lei e a sua execução de fato deve-se a diversos fatores, dentre eles, a escassez de pesquisadores na área da Dialetologia e, especialmente, as dificuldades de se percorrer a vasta extensão territorial do Brasil, principalmente nos idos de 1950.

Então, decorrido esse tempo e tendo também se desenvolvido as pesquisas dialetológicas que deram origem a seis atlas publicados no Brasil²⁸, a proposta de um atlas nacional tomou forma e deu origem ao Projeto Atlas Linguístico do Brasil (Projeto ALiB), que visa a

[...] documentar o país de Norte a Sul, de Leste a Oeste, e lembrando [Antenor] Nascentes, ‘do Oiapoque ao Chuí’: são 250 pontos a constituírem a sua rede de localidade que reunirão 1.100 informantes a serem documentados (CARDOSO, 2006, p. 29).

²⁸ Além do *Atlas Prévio dos Falares Baianos* (1963), foram publicados também o *Atlas Linguístico do Sergipe* (1973), o *Esboço de um Atlas Linguístico de Minas Gerais* (1977), o *Atlas Linguístico da Paraíba* (1984), o *Atlas Linguístico do Paraná* (1990) e o *Atlas Linguístico-Etnográfico da Região Sul do Brasil* (2002) (PROJETO ALiB. ATLAS NACIONAIS. Disponível em: <http://twiki.ufba.br/twiki/bin/view/Alib/AtlasNacionais>).

Isso com o objetivo mais amplo de documentar a variante brasileira do português em uso nas localidades pesquisadas, com vistas a mapear todo o território nacional e a gerar uma fotografia da língua tal qual se realiza. Dessa forma, o produto dessa pesquisa poderá fundamentar outros estudos na área, bem como auxiliar o próprio ensino da língua portuguesa, que poderá ser vista e conhecida em sua feição real.

Esse Projeto é de natureza pluridimensional, é focado na área urbana e é de alcance nacional, com sede na Bahia, coordenado por um comitê interinstitucional composto por 09 integrantes ligados a 09 Instituições de Ensino Superior brasileiras. Há a diretora presidente, Profa. Suzana Alice Marcelino da Silva Cardoso, da Universidade Federal da Bahia, a diretora executiva Profa. Jacyra Andrade Mota, também da UFBA e mais 07 diretores científicos (Abdelhak Razky – Universidade Federal do Pará; Maria do Socorro Silva de Aragão – Universidades Federais do Ceará e da Paraíba; Ana Paula Antunes Rocha – Universidade Federal de Ouro Preto; Vanderci de Andrade Aguilera – Universidade Estadual de Londrina; Aparecida Negri Isquerdo – Universidade Federal de Mato Grosso do Sul; Felício Wessling Margotti – Universidade Federal de Santa Catarina e Cléo Wilson Altenhofen – Universidade Federal do Rio Grande do Sul).

O Projeto ALiB faz uso de um questionário linguístico para a coleta de dados em uma rede de pontos composta por 250 localidades, tendo como informantes brasileiros naturais dessas localidades de acordo com um perfil previamente estipulado.

Quanto à rede de pontos, foram selecionadas, dentre as 250 localidades, capitais e localidades do interior, definidas de acordo com a proposta inicial de Nascentes, como mencionara Cardoso (2006) e também de acordo com a

importância sócio-histórico-geográfica dessas localidades. Em relação aos informantes, o Comitê de Coordenação Nacional do Projeto ALiB definiu o seguinte perfil: sexos masculino e feminino, de duas faixas etárias (18 a 30 e 50 a 65 anos), escolaridade: Ensino Fundamental (nas capitais e localidades do interior) e Ensino Superior (nas capitais), nascidos e criados na localidade pesquisada e de pais também provenientes dessa mesma região linguística. São inquiridos 04 informantes em cada localidade do interior e 08 em cada capital, para privilegiar, além dos aspectos diatópicos, dimensões diagenéricas, diageracionais e diastráticas.

Já em relação ao questionário, o instrumento de coleta contempla realizações fonético-fonológicas (159 perguntas); prosódicas (11 questões); semântico-lexicais – 14 áreas semânticas (202 perguntas); morfossintáticas (49 questões); pragmáticas (04 perguntas) e metalinguísticas (06 perguntas), além de quatro temas para discursos semidirigidos e um texto para leitura.

Segundo dados disponíveis no site do Projeto, já foram concluídos 93,6% das localidades a serem pesquisadas, com entrevistas com 1.036 dos 1.100 informantes, totalizando 80% dos Estados com inquéritos concluídos²⁹. Assim, faltam apenas 16 localidades das 250 do total, a serem visitadas (PROJETO ALiB, 2012. Disponível em <http://twiki.ufba.br/twiki/bin/view/Alib/WebHome>).

Dessa forma, o projeto vai caminhando para outras fases de sua execução, como a elaboração do primeiro volume do Atlas, que agregará informações relativas às capitais dos Estados, bem como evolui para caminhos antes não percorridos pelos atlas linguísticos produzidos no Brasil, como os estudos da prosódia, por exemplo.

²⁹ Dados consultados em 05 de novembro de 2012.

O Projeto ALiB congrega esforços de profissionais de diversas áreas, como dialetólogos e geolinguistas, mas também de cartógrafos, geógrafos, dentre outros, para que se torne conhecida e reconhecida a riqueza da variante brasileira do português, evidenciada pela variedade de usos e particularidades que têm se tornado visíveis graças aos estudos dos dados documentados pelo Projeto.

Este trabalho também se vincula ao Projeto ALiB, uma vez que tem como *corpus* dados inéditos desse Projeto, cuja utilização teve prévia autorização da presidente do Comitê Nacional de Coordenação do Projeto ALiB, que faculta o uso desses dados apenas para trabalhos orientados por membros do Comitê de Coordenação Nacional do ALiB. Dentre as 250 localidades que integram a rede de pontos do Projeto, este trabalho contempla 25 delas, as capitais brasileiras, perfazendo um total de 200 informantes (08 por capital). As áreas semânticas selecionadas para a pesquisa foram escolhidas por representarem aspectos nacionais, sendo também áreas propícias para registros de cunho regional, afetadas pelas condições sócio-históricas das localidades pesquisadas e também pelos contatos entre povos, como as *atividades agropastoris*, a *fauna* e a *alimentação e cozinha*.

Para melhor visualizar as áreas semânticas selecionadas, seguem três quadros contendo as perguntas do QSL cujas respostas foram estudadas, de acordo com a área semântica a que se vinculam:

| Atividades agropastoris | | |
|--------------------------------|--------------------|--|
| Pergunta/QSL | Lema | Formulação da pergunta |
| 39 | Tangerina/mexerica | “As frutas menores que a laranja, que se descascam com a mão, e, normalmente, deixam um cheiro na mão? Como elas são?” |
| 40 | Amendoim | “O grão coberto por uma casquinha dura, que se come assado, cozido, torrado ou moído?” |

| | | |
|----|--|--|
| 41 | Camomila | “Umás florezinhas brancas com miolo amarelinho, ou florezinhas secas que se compram na farmácia ou no supermercado e servem para fazer um chá amarelinho, cheiroso, bom para dor de barriga de nenê/bebê e até de adulto e também para acalmar?” |
| 42 | Penca | “Cada parte que se corta do cacho da bananeiro para pôr para madurar/amadurecer?” |
| 43 | Banana dupla/felipe/gêmeas | “Duas bananas que nascem grudadas?” |
| 44 | Parte terminal da inflorescência da bananeira/umbigo/coração | “A ponta roxa do cacho da banana?” |
| 45 | Espiga | “Quando se vai colher o milho, o que é que se tira do pé?” |
| 46 | Sabugo | “Quando se tira da ____ (cf. item 45) todos os grãos do milho, o que sobra?” |
| 47 | Soca/touceira | “Depois que se corta o pé de arroz ou de fumo, ainda fica uma pequena parte no chão. Como se chama essa parte?” |
| 48 | Girassol | “Flor grande, amarela, redonda, com uma rodela de sementes no meio?” |
| 49 | Vagem do feijão/bainha | “Onde é que ficam os grãos do feijão, no pé, antes de serem colhidos?” |
| 50 | Mandioca/aipim | “Aquele raiz branca por dentro, coberta por uma casca marrom, que se cozinha para comer?” |
| 51 | Mandioca | “Uma raiz parecida com ____ (cf. item 50) que não serve para comer e se rala para fazer farinha (polvilho, goma)?” |
| 52 | Carrinho de mão/carriola | “Um veículo de uma roda, empurrado por uma pessoa, para pequenas cargas em trechos curtos?” |
| 53 | Hastes do carrinho de mão | “As duas partes em que a pessoa segura para empurrar o (a) ____ (cf. item 52)?” |
| 54 | Cangalha/forquilha | “A armação de madeira, que se coloca no pescoço dos animais (porco, terneiro/bezerro, |

| | | |
|----|--------------------------------------|---|
| | | carneiro, vaca), para não atravessarem a cerca?” |
| 55 | Cangalha | “A armação de madeira que se coloca no lombo do cavalo ou do burro para levar cestos ou cargas?” |
| 56 | Canga | “A peça de madeira que vai no pescoço do boi, para puxar o carro ou o arado?” |
| 57 | Jacá/balaio | “Aqueles objetos de vime, de taquara, de cipós trançados, para levar batatas (mandioca, macaxeira, aipim, etc.), no lombo do cavalo ou do burro?” |
| 58 | Bolsa/bruaca | “E quando se usam objetos de couro, com tampa, para levar farinha, no lombo do cavalo ou do burro?” |
| 59 | Borrego (do nascer até...) | “A cria da ovelha logo que nasce? E até que idade se dá esse nome?” |
| 60 | Perda da cria | “Como se diz quando a fêmea de um animal perde a cria?” |
| 61 | Trabalhador de enxada em roça alheia | “O homem que é contratado para trabalhar na roça de outro, que recebe por dia de trabalho?” |
| 62 | Picada/atalho estreito | “O que é que se abre com facão, a foice para passar por um mato fechado?” |
| 63 | Trilho/caminho/vereda/trilha | “O caminho, no pasto, onde não cresce mais grama, de tanto o animal ou o homem passarem por ali?” |

Quadro I – Questionário Semântico-lexical/ALiB, área semântica *atividades agropastoris*

| Fauna | | |
|---------------------|------------------------------|---|
| Pergunta/QSL | Lema | Formulação da pergunta |
| 64 | Urubu | “A ave preta que come animal morto, podre?” |
| 65 | Colibri/beija-flor | “O passarinho bem pequeno, que bate muito rápido as asas, tem o bico comprido e fica parado no ar?” |
| 66 | João-de-barro | “A ave que faz a casa com terra, nos postes, nas árvores e até nos cantos da casa?” |
| 67 | Galinha d’angola/guiné/cocar | “A ave de criação parecida com a galinha, de penas pretas com pintinhas brancas?” |

| | | |
|----|----------------------------|--|
| 68 | Papagaio | “A ave de penas coloridas que, quando presa, pode aprender a falar?” |
| 69 | Sura | “Uma galinha sem rabo?” |
| 70 | Cotó | “Um cachorro de rabo cortado?” |
| 71 | Gambá | “O bicho que solta um cheiro ruim quando se sente ameaçado?” |
| 72 | Patas dianteiras do cavalo | “As patas dianteiras do cavalo?” |
| 73 | Crina do pescoço | “O cabelo em cima do pescoço do cavalo?” |
| 74 | Crina da cauda | “O cabelo comprido na traseira do cavalo?” |
| 75 | Lombo | “A parte do cavalo onde vai a sela?” |
| 76 | Anca/garupa/cadeira | “A parte larga atrás do ____ (cf. item 75)?” |
| 77 | Chifre | “O que o boi tem na cabeça?” |
| 78 | Boi sem chifre | “O boi sem ____ (cf. item 77)?” |
| 79 | Cabra sem chifre | “A cabra que não tem ____ (cf. item 77)?” |
| 80 | Úbere | “Em que parte da vaca fica o leite?” |
| 81 | Rabo | “A parte com que o boi espanta as moscas?” |
| 82 | Manco | “O animal que tem uma perna mais curta e que puxa de uma perna?” |
| 83 | Mosca varejeira | “Um tipo de mosca grande, esverdeada, que faz um barulhão quando voa?” |
| 84 | Sanguessuga | “Um bichinho que se gruda nas pernas das pessoas quando elas entram num córrego ou banhado?” |
| 85 | Libélula | “O inseto de corpo comprido e fino, com quatro asas bem transparentes, que voa e bate a parte traseira na água?” |
| 86 | Bicho de fruta | “Aquele bichinho branco, enrugadinho, que dá em goiaba, em coco?” |
| 87 | Coró | “Aquele bichinho que dá em esterco, em pau podre?” |
| 88 | Pernilongo | “Aquele inseto pequeno, de perninhas compridas, que canta no ouvido das pessoas de noite?” |

Quadro II – Questionário Semântico-lexical/ALiB, área semântica *fauna*

| Alimentação e cozinha | | |
|------------------------------|--------------------------|--|
| Pergunta/QSL | Lema | Formulação da pergunta |
| 176 | Café da manhã | “Como se chama a primeira refeição do dia, feita pela manhã?” |
| 177 | Geleia | “Como se chama a pasta feita de frutas para passar no pão, biscoito?” |
| 178 | Carne moída | “Como se chama a carne depois de triturada na máquina?” |
| 179 | Curau / canjica | “Como se chama uma papa cremosa feita com coco e milho verde ralado, polvilhada com canela?” |
| 180 | Curau | “E essa mesma papa, com milho verde ralado, sem coco, como é que chama?” |
| 181 | Mungunzá / canjica | “Como se chama aquele alimento feito com grãos de milho branco, coco e canela?” |
| 182 | Aguardente | “Como se chama a bebida alcóolica feita de cana-de-açúcar?” |
| 183 | Empanturrado | “Quando uma pessoa acha que comeu demais, ela diz: Comi tanto que estou...?” |
| 184 | Glutão | “Como se chama uma pessoa que normalmente come demais?” |
| 185 | Bala / confeito / bombom | “Como se chama aquilo embrulhado em papel colorido que se chupa?” |
| 186 | Pão francês | “Como se chama isto? (mostrar)” |
| 187 | Pão bengala | “Como se chama isto? (mostrar)” |

Quadro III – Questionário Semântico-lexical/ALiB, área semântica *alimentação e cozinha*

Assim, a partir da coleta das respostas fornecidas para as perguntas elencadas, as unidades lexicais auferidas foram lançadas no *software* SGP – Sistema Geral de Pesquisa, que auxiliou este trabalho como ferramenta no trato das informações e que é apresentado a seguir.

3.2. O SGP – SISTEMA GERAL DE PESQUISA: A TECNOLOGIA DA INFORMAÇÃO AUXILIANDO A PESQUISA LINGUÍSTICA

Tendo em vista o vulto do *corpus* analisado, este trabalho buscou ajuda da Tecnologia da Informação para auxiliar no manejo das informações, com vistas a permitir uma melhor manipulação dos dados e, dessa forma, possibilitar também uma gama maior de condições de análise, com a facilitação do acesso às informações, a partir de diferentes pontos de vista.

Assim, buscou-se o trabalho de um analista de sistemas, profissional que trabalha com o desenvolvimento de sistemas de informação, para que fosse produzido um *software* específico para armazenamento e manipulação dos dados.

Após reuniões iniciais e posteriores testes, foi desenvolvido o Sistema Geral de Pesquisa, *software* especialmente produzido para esta dissertação, mas que pode ser utilizado em outros tipos de pesquisa do gênero, tendo em vista possibilitar diversos cadastros e consequentes relatórios.

O referido programa foi desenvolvido na Linguagem Java de programação, utilizando banco de dados Postgre e fazendo uso de Framework do Governo Federal, o Deimoiselle. Essa configuração técnica possibilita que o sistema seja disponibilizado em servidor *web*, o que faculta o acesso, via *internet*, em qualquer lugar e por qualquer computador, além de poder ser utilizado por mais de um usuário, simultaneamente.

Na sequência, esse sistema é apresentado e caracterizado por partes. Esse procedimento, além de demonstrar a forma de armazenamento dos dados deste estudo, permite a divulgação da ferramenta tecnológica adotada, que poderá ser usada futuramente em outras pesquisas.

O SGP prima pela interatividade com o usuário, além da facilidade de manuseio, não havendo a necessidade de extensos treinamentos ou mesmo conhecimento aprofundado em informática para sua manipulação. O sistema requer apenas um servidor para armazenar o banco de dados que, por sua vez, fornecerá as informações disponibilizadas por meio de diferentes tipos de relatórios.

A tela inicial já apresenta todos os cadastros possíveis, havendo também condições de serem cadastrados outros projetos de pesquisa ou outras pesquisas para um mesmo usuário, como pode ser visto na Figura 1, a seguir:

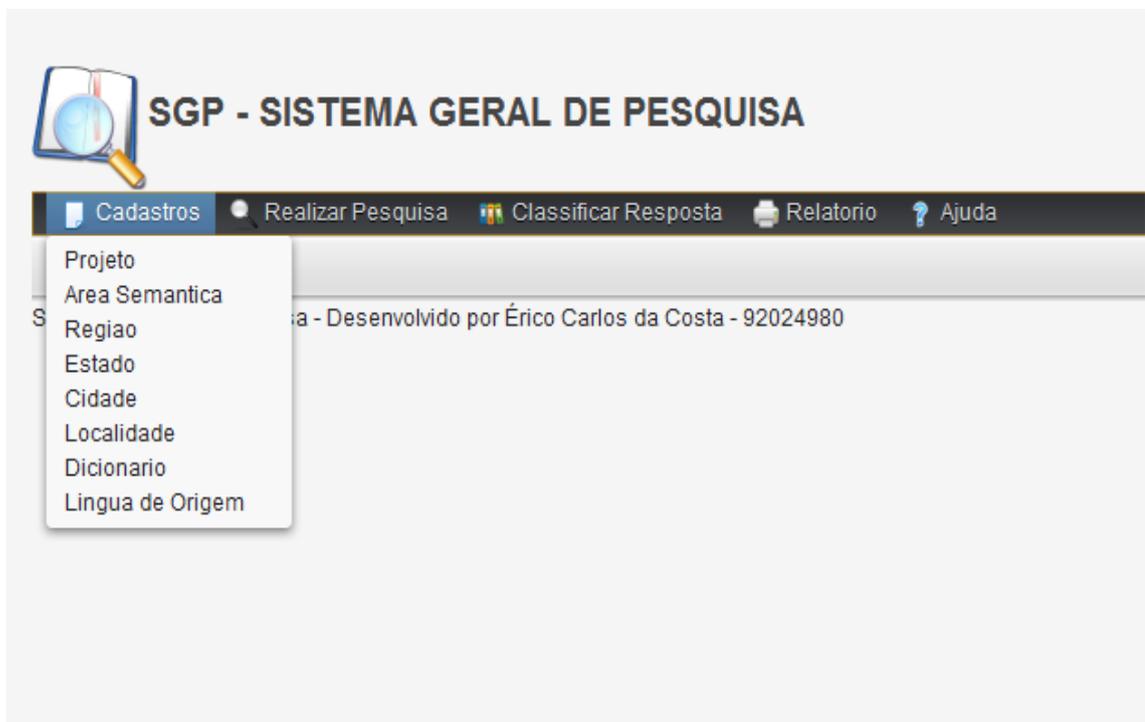


Figura 1: Tela inicial do SGP

Cadastrados os dados do Projeto, áreas semânticas e as informações acerca das localidades e dos informantes³⁰, a continuidade da primeira etapa do

³⁰ Por ser um dos critérios da pesquisa geolinguística a amostragem qualitativa dos dados, não é importante a identificação pessoal dos entrevistados. Sendo assim, os dados dos informantes armazenados pelo sistema seguem a codificação adotada pelo Projeto ALiB, formada pelo número da localidade seguido do número do informante que, por sua vez, sintetiza o perfil desse informante. Nesse sentido, os números ímpares representam o sexo masculino (1, 3, 5 e 7) e os pares, o feminino (2, 4, 6 e 8). Os números baixos representam informantes do

trabalho deu-se com o cadastramento das perguntas relacionadas para a pesquisa que aparecem na tela apresentada a seguir:

The screenshot shows the 'SISTEMA DE PESQUISA' interface. At the top, there are navigation links for 'Cadastros' and 'Ajuda'. Below that, a 'Pergunta' tab is active. The form includes a dropdown for 'Area Semantica' (set to 'Atividades agropastoris'), a text area for 'Cabeça da Pergunta' containing the question text, a 'Numero da Pergunta' field with the value '39', and a 'Resposta Esperada' field with the value 'tangerina / mexerica'. There are three buttons: 'Gravar', 'Excluir', and 'Limpar'.

| AREA SEMANTICA | CABECA DA PERGUNTA | ID | NUMERO DA PERGUNTA | RESPOSTA ESPERA |
|-------------------------|--|----|--------------------|------------------|
| Atividades agropastoris | Como se chamam as frutas menores que a laranja, que se descascam com a mão e, normalmente, deixam um cheiro na mão? Como elas são? | 1 | 39 | tangerina / mex |
| Atividades agropastoris | Como se chama o grão coberto por uma casquinha dura, que se come assado, cozido, torrado ou moído? | 2 | 40 | amendoim |
| Atividades agropastoris | Como se chama cada parte que se corta do cacho da bananeira para pôr para madurar/amadurecer? | 10 | 42 | penca |
| Atividades agropastoris | Como se chamam duas bananas que nascem grudadas? | 11 | 43 | banana dupla/fe |
| Atividades agropastoris | Como se chama a ponta roxa do cacho da banana? | 12 | 44 | parte terminal d |
| Atividades agropastoris | Quando se vai colher o milho, o que é que se tira do pé? [Quando se vai à feira comprar milho, compra-se o quê?] | 13 | 45 | espiga |
| Atividades agropastoris | Depois que se corta o pé de arroz ou de fumo, ainda fica uma pequena parte no chão. Como se chama essa parte? | 15 | 47 | soca/touceira |

Figura 2: Cadastro de perguntas no SGP

Como se pode verificar, no cadastramento da pergunta, já se atribui também a ela a área semântica pertinente, o ID (identificação dentro do sistema), bem como seu número de identificação – número no QSL/ALiB – e a cabeça da pergunta (que apresenta potenciais respostas), também fornecida pelo questionário do referido Projeto. E, visando também à facilitação na manipulação das informações, o SGP visualiza, na mesma tela, todas as perguntas já cadastradas para aquela área semântica, abrindo a possibilidade de gravação, edição e exclusão das informações cadastradas.

Ensino Fundamental (de 1 a 4) e os mais altos, egressos do Ensino Superior (de 5 a 8). Para a posterior análise, caso seja necessário, aí sim serão consultados dados pertinentes aos informantes, como histórico de vida e informações pessoais.

Após a alimentação das perguntas, inserem-se as respostas fornecidas pelos informantes³¹, como se pode visualizar na tela a seguir, representada pela

Figura 3:

| AREA SEMANTICA | ID | INFORMANTE |
|-------------------------|----|------------|
| Atividades agropastoris | 2 | 002-3 |
| Atividades agropastoris | 1 | 002-1 |
| Atividades agropastoris | 3 | 002-2 |
| Fauna | 4 | 002-2 |
| Alimentação | 7 | 002-2 |

| AREASEMANTICA | CABECAPERGUNTA | ID | NUMEROPERGUNTA | RESPOSTAESPERADA |
|---------------|---|----|----------------|--------------------|
| Fauna | Como se chama a ave preta que come animal morto, podre? | 45 | 64 | urubu |
| Fauna | Como se chama o passarinho bem pequeno, que bate muito rápido as asas, tem o bico comprido e fica parado no ar? | 46 | 65 | colibri/beija-flor |

Figura 3: Cadastramento inicial no SGP

Assim, essa tela possibilita a posterior associação entre a resposta citada pelo informante³² e o respectivo perfil:

³¹ Como em todos os demais relatórios, o sistema permite o cruzamento das informações cadastradas, podendo-se recuperar dados inseridos em todas as abas e compará-los analiticamente, por meio das listagens emitidas pelo SGP.

³² Todos os informantes entrevistados pelo Projeto ALiB têm seus dados pessoais armazenados numa ficha, de modo a ser possível resgatar essas e outras informações a partir desse formulário. Todavia, por questões éticas, todos esses dados são sigilosos, para garantir a impessoalidade da pesquisa, o que autoriza que sejam utilizadas as informações acerca de profissão, idade e experiência de vida, mas nunca dados pessoais (nomes, endereços ou telefones).

Figura 4: Cadastro de resposta no sistema SGP

Aqui, há a possibilidade de duas formas de cadastramento da resposta, seja pela aba posterior, abaixo do *grid* inicial, ou mesmo pela seleção na aba **pergunta**, condição essa que auxilia a agilidade no cadastro dos dados.

Também nessa aba podem ser cadastradas todas as respostas fornecidas pelos informantes, haja vista que a ordem proferida pode revelar aspectos relevantes para a interpretação dos dados.

Os próximos cadastros dizem respeito aos dados necessários para subsidiar a análise: dicionários pesquisados e língua de origem das respostas:

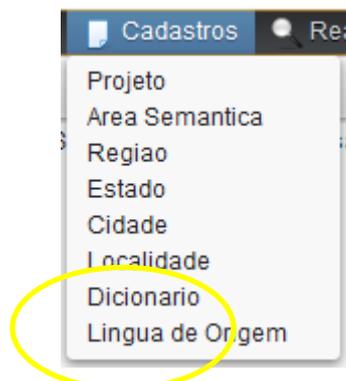


Figura 5: Tela de cadastros do SGP

Depois de cadastradas todas as informações, passa-se ao tratamento dos dados, momento em que foi realizada a classificação das respostas arquivadas, acrescentando-se informações acerca da língua de origem, do étimo e das definições lexicográficas, conforme ilustra a Figura 6, a seguir:

Étimo

Lingua de Origem

Resposta

| DESCRICAÇÃO | ID | LINGUA DE ORIGEM | LISTADICIONARIO | RESPOSTA |
|-------------|----|------------------------|-----------------|------------------|
| | 2 | LP - Língua Portuguesa | ☐ | beija-flor |
| | 3 | LE - Língua Espanhola | ☐ | raposa |
| | 4 | LE - Língua Espanhola | ☐ | chifre |
| | 5 | LT - Língua Tupi | ☐ | carapanã |
| | 6 | LP - Língua Portuguesa | ☐ | pão casca grossa |
| | 7 | LP - Língua Portuguesa | ☐ | pão comum |

Definições

Figura 6: Tela de classificação de respostas do SGP

Nessa tela, de forma intuitiva e prática, como em todo o sistema, também são visualizadas as respostas já classificadas, segundo a língua de origem.

Cadastradas todas as informações, o sistema permite a geração de relatórios com base em qualquer um dos campos alimentados, inclusive com as respostas classificadas, entrecruzando as informações e fornecendo subsídios para a análise, como percentuais de ocorrência, áreas de registro dos itens lexicais e informações detalhadas acerca do conjunto das unidades léxicas cadastradas, dentre outros.

Dentre os relatórios gerados pelo sistema, pode-se destacar um daqueles que forneceram dados para a análise léxico-semântica, o que seleciona as informações a partir da língua de origem cadastrada das unidades lexicais que constituem o *corpus* do estudo, o que permite gerar as listagens do tipo da visualizada na Figura 7:

| LISTAGEM DE RESPOSTAS COM ÁREA SEMÂNTICA POR LINGUA DE ORIGEM | |
|---|-------------------|
| Área Semântica | Resposta |
| Gótico | |
| Alimentação | roupa velha |
| Qtd. Respostas por lingua de origem : 1 | |
| LA - Línguas Africanas | |
| Alimentação | canjica |
| Alimentação | munguzá |
| Alimentação | canjica sem leite |
| Alimentação | munguzá |
| Alimentação | canjicada |
| Alimentação | canjica branca |
| Alimentação | canjica de milho |
| Alimentação | canjição |
| Alimentação | angu |
| Alimentação | xerém |
| Qtd. Respostas por lingua de origem : 10 | |
| LAI - Língua Alemã | |
| Alimentação | chimia |
| Alimentação | guisado |
| Qtd. Respostas por lingua de origem : 2 | |

Figura 7: Listagem de respostas por língua de origem

E também se pode citar o relatório que fornece informações de natureza diatópica das unidades lexicais cadastradas no sistema. Nesse caso, pode-se observar tanto o informante que mencionou a resposta como essa segundo a pergunta geradora, recurso que abriu o leque de possibilidades de estudos dos dados:

| RELATÓRIO DE RESPOSTAS POR INFORMANTE E REGIÃO | | | | |
|---|--------|---|------|-----------------|
| N. Local | Inform | pergunta | id | resposta |
| Alimentação | | | | |
| Centro-Oeste - Goiás - Goiânia | | | | |
| 123 | 2 | 179 - Como se chama uma papa cremosa feita com coco e milho verde ralado, polvilhada com canela? | 107 | mingau de milho |
| 123 | 2 | 180 - E essa mesma papa, com milho verde ralado, sem coco, como é que se chama? Pedir para descrever como se faz. | 107 | mingau de milho |
| Quantidade por região : 2 | | | | |
| Centro-Oeste - Mato Grosso do Sul - Campo Grande | | | | |
| 115 | 2 | 179 - Como se chama uma papa cremosa feita com coco e milho verde ralado, polvilhada com canela? | 107 | mingau de milho |
| Quantidade por região : 1 | | | | |
| Nordeste - Alagoas - Maceió | | | | |
| 077 | 3 | 182 - Como se chama a bebida alcoólica feita de cana-de-açúcar? | 2327 | pitu |
| Quantidade por região : 1 | | | | |
| Nordeste - Bahia - Salvador | | | | |
| 093 | 3 | 179 - Como se chama uma papa cremosa feita com coco e milho verde ralado, polvilhada com canela? | 107 | mingau de milho |
| 093 | 4 | 182 - Como se chama a bebida alcoólica feita de cana-de-açúcar? | 2476 | jacaré |
| Quantidade por região : 2 | | | | |

Figura 8: Listagem diatópica de indigenismos segundo a capital brasileira

Esses, bem como os demais relatórios gerados pelo SGP, possibilitam dimensionar o universo estudado, bem como sintetizam as informações desejadas, o que auxilia o pesquisador, uma vez que se podem visualizar dados parciais que interessam nos diversos momentos e perspectivas da análise dos dados.

Assim, a análise apresentada no capítulo a seguir foi viabilizada a partir da adoção do SGP, bem como dos procedimentos metodológicos discutidos neste capítulo.

3.3 TRATAMENTO DOS DADOS

Tendo em vista a complexidade dos dados analisados, foi necessário sistematizar o tratamento das informações, de modo que as análises pudessem contemplar as diversas possibilidades oferecidas pelo estudo das unidades lexicais.

Assim, realizado o primeiro passo, que foi a alimentação dos dados no sistema SGP, anteriormente apresentado, este estudo buscou expressar, de acordo com as ciências que lhe deram subsídios, apresentadas no Capítulo II, e segundo a

metodologia aqui apresentada, vários prismas de investigação, consoante o possibilitado pelo *corpus*.

Nesse sentido, as discussões apresentadas foram orientadas pela análise que buscou duas vertentes, uma da macro e uma da microestrutura, isto é, o Capítulo IV, que apresenta e discute os dados, foi organizado a partir, inicialmente, do macrouniverso dos dados, ou seja, apresenta e analisa o conjunto geral das informações, todas as unidades léxicas de origem indígena registradas. Depois, num segundo momento, os dados foram separados segundo a área semântica e organizados em conformidade com a ordem em que são apresentadas no Questionário Semântico-Lexical do Projeto ALiB.

E os estudos realizados têm dois vieses, um léxico-semântico e um diatópico. Além disso, a análise foi ilustrada com elementos gráficos, para que se pudesse fornecer uma melhor visualização e, por conseguinte, uma melhor compreensão dos dados.

Para isso, foram confeccionados gráficos, quadros e cartas linguísticas, apresentados no decorrer do trabalho, que ora apontam para a dimensão léxico-semântica, ora para a distribuição diatópica dos itens léxicos e suas respectivas características.

Os gráficos visualizaram a produtividade das unidades léxicas frente à base linguística, à área semântica em que foram documentadas e à sua participação no universo geral dos dados. Os quadros, por sua vez, em maior número nesta Dissertação, trouxeram diversas informações, como a ocorrência geral dos itens lexicais numa dada área semântica, traços léxico-semânticos comparativos entre unidades lexicais, assim como outros dados relevantes para o estudo.

No que diz respeito às cartas linguísticas, há informações que devem ser aqui destacadas, haja vista tratarem-se de um recurso que requer vários momentos para sua confecção. Inicialmente, deve-se apontar que a base cartográfica utilizada por este trabalho para a preparação desses mapas é advindo do Projeto Atlas Linguístico do Brasil, elaborada pela geógrafa e cartógrafa Ana Regina Telles. Algumas delas foram trabalhadas por Luciene Freitas-Marins, Profa. Mestre da Universidade Estadual de Mato Grosso do Sul, e outras pela autora deste trabalho.

Os dados analisados geraram cartas de dois gêneros: geral e específico. As de natureza geral contemplaram aspectos diatópicos, por exemplo, enquanto as específicas trouxeram outras vertentes de análise, como apresentando a produtividade e a base linguística, por exemplo. Dentre essas, ora foram focalizadas as áreas semânticas na totalidade, ora cada qual, em separado. Algumas cartas trouxeram ainda informações acerca de uma pergunta em específico ou, nas mais particularizadas, foram visualizados aspectos de uma única unidade léxica.

Dessa maneira, pretendeu-se dar visibilidade aos dados e os apresentar de forma que o leitor deste trabalho possa associar o texto às ilustrações, no sentido de que um complete o outro, haja vista que constituem, em conjunto, o Capítulo IV, que versa sobre a análise do *corpus* pesquisado.

E realizadas, pois, as considerações acerca dos procedimentos metodológicos adotados, a continuidade do trabalho dá-se, pois, com a análise dos dados, apresentada pelo Capítulo IV, a seguir.

CAPÍTULO IV

APRESENTAÇÃO E ANÁLISE DOS DADOS

Sendo o homem um ser social, sua língua também revela olhares desse homem sobre seu mundo e sua realidade, além de transmitir conhecimentos, ideias e tradições.

No decorrer deste trabalho, discutiu-se que, além de a língua ser um produto social, assim também o léxico transmite, mais que fatores linguísticos, mas ainda determinantes socioculturais, históricos e geográficos, por exemplo. E, no Brasil, o português adquiriu cor local, dentre outros fatores, dada a inter-relação entre os povos que aqui se encontraram, como os europeus e os indígenas, relação essa que é abordada por este estudo no tocante à presença de indigenismos no repertório vocabular do homem urbano habitante das capitais brasileiras.

Dessa maneira, este Capítulo apresenta o estudo de unidades léxicas dessa categoria sob a luz do aporte teórico apresentado no Capítulo II e de acordo com a Metodologia discutida no Capítulo III. E, como já mencionado, a análise realizada teve os enfoques léxico-semântico e diatópico, com vistas a atender aos objetivos propostos na Introdução deste trabalho.

Para isso, este capítulo foi dividido em duas partes principais, a saber: de início, analisou-se o conjunto dos dados, de forma geral, e depois, os subtópicos seguintes focalizaram cada área semântica em separado, de maneira a aprofundar as análises, especialmente com o viés léxico-semântico.

Sendo assim, a partir das respostas fornecidas pelos informantes das 25 capitais brasileiras em estudo, para as 62 perguntas relacionadas às três áreas semânticas contempladas por este estudo – *atividades agropastoris*, *fauna* e *alimentação e cozinha*, foi auferido um total de 800 itens lexicais³³, assim distribuídos: 305 (*atividades agropastoris*), 308 (*fauna*) e 187 (*alimentação e cozinha*). Esses dados foram analisados na sequência deste Capítulo, tendo-se reservado a parte inicial para o cotejo do montante geral das respostas, segundo diferentes olhares.

No conjunto dos itens lexicais catalogados, registraram-se 59 indigenismos, o que equivale a 7,2% das respostas documentadas. A maioria das respostas foi de origem portuguesa, que resultaram em 500 unidades léxicas, havendo também a ocorrência de itens lexicais de outras bases linguísticas³⁴, representadas no Gráfico I, a seguir:

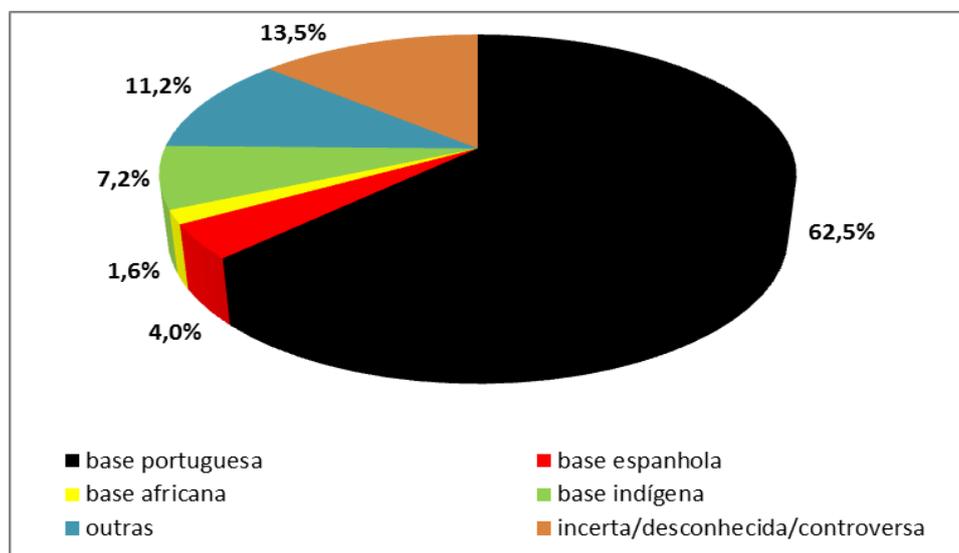


Gráfico I – Distribuição do total dos itens lexicais documentados de acordo com a base linguística

³³ De fato, houve 1.822 ocorrências. O montante de 800 itens lexicais foi calculado a partir do agrupamento de variantes fonéticas e/ou morfológicas, que não vigoram individualmente neste trabalho tendo em vista sua vertente lexical.

³⁴ Na legenda, o item lexical **outras** agrupada as unidades lexicais de origem alemã, árabe, gótica, francesa, inglesa, italiana e catalã, além das que têm origem toponímica e onomatopaica. Já o item lexical **incerta, desconhecida ou controversa** reuniu as unidades léxicas cuja etimologia não foi encontrada em nenhuma das obras pesquisadas.

Vê-se, a partir da análise do Gráfico I, que a língua do colonizador que substituiu a Língua Geral desde os tempos coloniais se sobrepôs a todas as demais línguas, incluindo-se as línguas indígenas. Além disso, percebe-se que muito há ainda para se estudar em termos etimológicos, haja vista haver 107 unidades léxicas (13%), cuja base etimológica não foi identificada nas obras de referência consultadas.

No que diz respeito aos dados de origem indígena, todas as unidades léxicas aqui registradas são de base tupi³⁵. Nota-se que a história linguística brasileira, no que tange às bases indígenas na formação do português do Brasil, é representada pela quantidade de tupinismos auferidos pelo presente trabalho. Reforça também essa perspectiva a afirmação de Cardoso que, ao estudar as línguas indígenas, registrou que “a grande corrente léxica dos linguajares indígenas americanos veio, indiscutivelmente, do tupi” (CARDOSO, 1961, p. 36), o que se confirma pelos dados deste estudo.

Os 59 registros indígenas aqui estudados vinculam-se, em sua maioria, à área semântica da *fauna* (30 ocorrências), seguido pela das *atividades agropastoris* (21 registros) e com a menor produtividade na área da *alimentação e cozinha* (08 ocorrências)³⁶.

Esses dados aparentemente demonstram que a área da *fauna* lidera os registros, seguido de perto pela área das *atividades agropastoris*. Todavia, se considerada a relação entre o número de perguntas de cada área semântica e o percentual de ocorrências, nota-se um equilíbrio em termos de produtividade de

³⁵ Foi documentada também *guampa*, que, de acordo com Machado (1987), seria de origem *mapuche*, língua ameríndia não falada no Brasil, mas no Chile; todavia, esse mesmo autor acredita que o registro em língua portuguesa dessa unidade lexical seria advindo do espanhol platino, origem essa considerada por este trabalho.

³⁶ As ocorrências dos indigenismos em cada área semântica foram detalhadamente analisadas nos subtópicos a elas pertinentes, na sequência deste capítulo.

indigenismos: enquanto a área da *alimentação e cozinha* reúne apenas 12 questões, a da *fauna* e a das *atividades agropastoris* contêm 25 questões cada uma. Nesse sentido, se considerarmos a relação entre o número de questões e o total de ocorrências por área semântica, teríamos percentuais semelhantes de registros: 52% para a *fauna*, 45,6% para as *atividades agropastoris* e 40% para *alimentação e cozinha*³⁷.

A forte presença de itens lexicais de origem indígena nessas áreas semânticas que congregam um maior número de elementos tipicamente nacionais, como os animais, atividades de campo e hábitos alimentares pode ser explicada pelo próprio contexto do Brasil Colônia em que os indígenas dominavam a realidade local e, conseqüentemente, imprimiram suas próprias tecnologias que foram assimiladas pelos europeus inexperientes no trato com a terra e com os novos hábitos que estavam se fixando na cultura brasileira.

Dessa forma, mesmo assim, se for considerada a média de questões, a área da *alimentação e cozinha* ainda revela um menor índice de indigenismos, haja vista que, mesmo tendo quase a metade da quantidade de perguntas das demais áreas semânticas, o registro de ocorrências aproximou-se da terça parte das *atividades agropastoris* e da quarta parte de *fauna*.

O cotejo entre o número de indigenismos e a respectiva área semântica é representado pelo Gráfico II, a seguir, que demonstra a dispersão dos itens lexicais em estudo segundo a área a que se vinculam:

³⁷ Cf. Quadros I, II e III (páginas 87, 89 e 91, respectivamente).

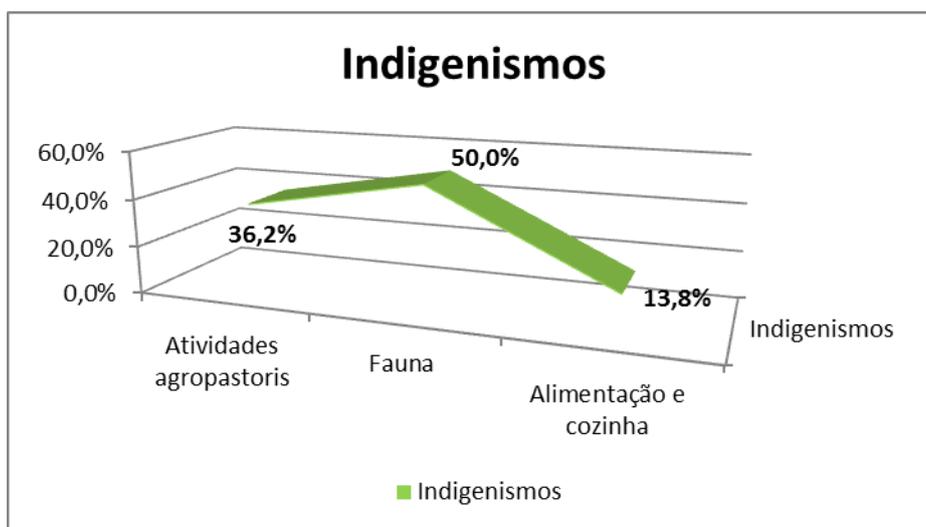


Gráfico II – Indigenismos

Em se tratando da produtividade das unidades lexicais em relação à distribuição diatópica, a região Norte lidera os registros, com um total de 38 indigenismos proferidos pelos 48 informantes entrevistados. Logo na sequência, registram-se 33 ocorrências no Nordeste, 20 no Centro-Oeste, 19 no Sudeste e 16 no Sul.

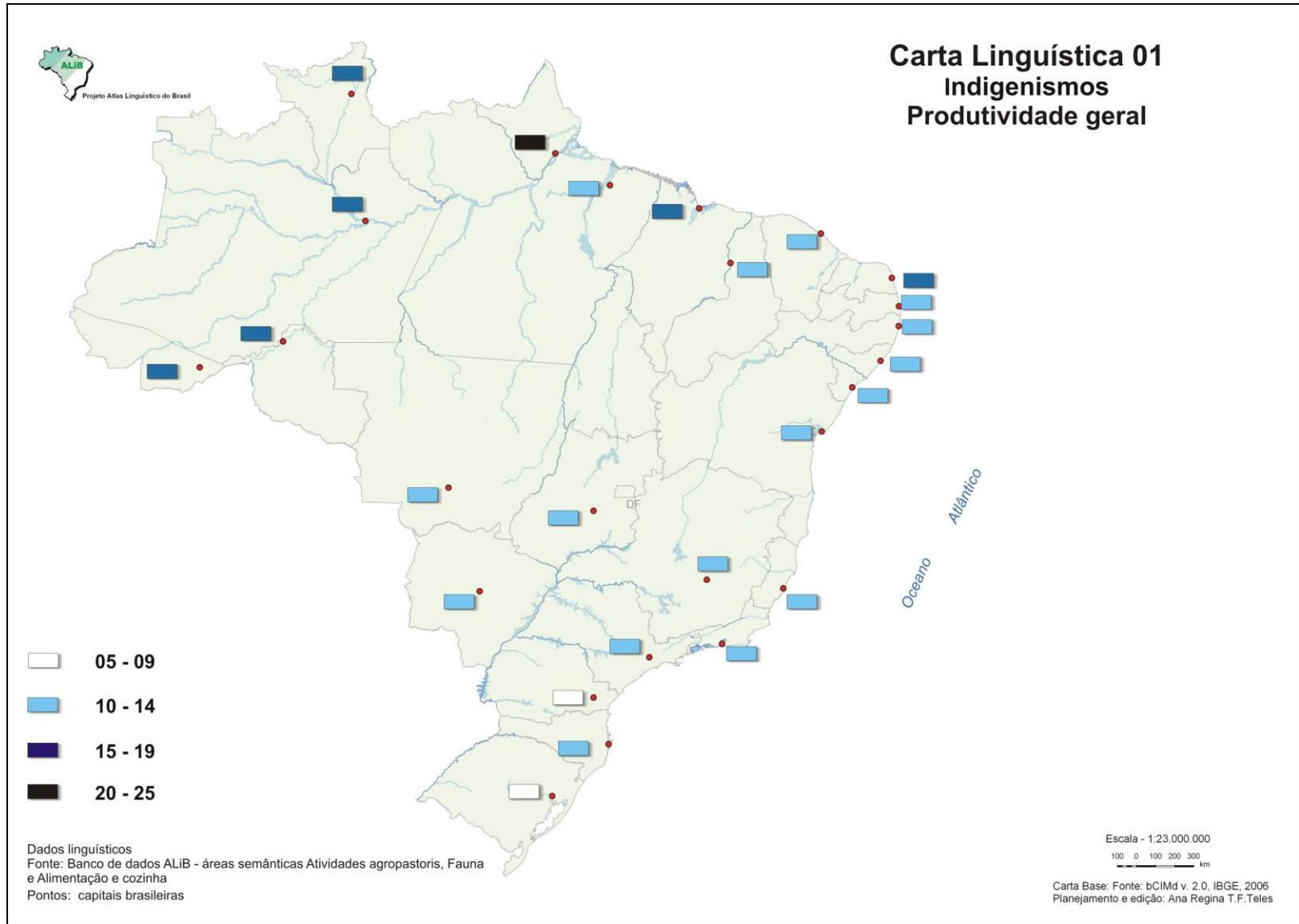
Vê-se assim que a região Norte, que congrega a maior população indígena no Brasil (IBGE, 2010), ainda tem marcada essa presença indígena no léxico de seus habitantes. Além disso, naquela região, é língua corrente em algumas localidades (e co-oficial em São Gabriel da Cachoeira/AM) o Nheengatu, língua derivada da Língua Geral Amazônica, advinda também do tupi antigo. Esse panorama pode ter contribuído para o alto registro de indigenismos nessa região, haja vista que essa parte do território nacional já era altamente povoada por nações indígenas, indicada “pelos muitos etnônimos registrados ao longo dos séculos XVII, XVIII e XIX” (MENÉNDEZ, 1992, p. 281). Nas proximidades da região conhecida como Madeira-Tapajós, no Amapá, inclusive, foi registrado o maior número de indigenismos: 21 registros em Macapá.

Já na região Nordeste, a segunda em registro de tupinismos, pode-se associar os 33 registros ao histórico dos indígenas naquela região, que comportou aldeamentos missionários especialmente durante o século XVIII (DANTAS; SAMPAIO; CARVALHO, 1992, p. 445-446) e ainda tem como característica o conservadorismo linguístico, em que diversas unidades léxicas em uso em épocas pretéritas mantêm-se na língua viva de seus habitantes, como a continuidade deste trabalho demonstrará. Além disso, esta é a região com o maior número de capitais contempladas pela rede de pontos do Projeto ALiB (09), o que congrega um total de 72 informantes entrevistados, do montante geral de 200 entrevistados considerados por esta pesquisa.

A terceira região, em registros, foi a Centro-Oeste, seguida pelo Sudeste e, por último, pelo Sul do Brasil. Os índices de produtividade no âmbito dessas regiões foram semelhantes, o que revela que a Língua Geral, por questões históricas, não se disseminou por esses territórios, como ocorreu no Norte e no Nordeste. Isso pode ser explicado pela colonização tardia do Centro-Oeste brasileiro, berço também de outras etnias, notadamente a partir do século XX, mais fortemente, e também pelo fato de o Sudeste, marcadamente, comportar metrópoles mundiais que recebem influências externas (como São Paulo e Rio de Janeiro) e também pela intensa presença africana nos tempos coloniais. Já o Sul tem nas suas bases étnicas especialmente povos europeus em detrimento de indígenas, que colonizaram aquela região e lhe imprimiram características singulares que ainda perduram nos dias atuais.

Essa distribuição diatópica pode ser resumida a partir da Carta Linguística I, que congrega o índice de ocorrências de indigenismos por capital brasileira e é apresentada na sequência.

Os dados mapeados demonstram claramente a alta concentração de indigenismos no Norte, diminuindo gradativamente a partir do Nordeste, Centro-Oeste e Sudeste, até um quase apagamento no Sul do Brasil. Vê-se assim que os processos de colonização do território deixaram marcas no léxico dos habitantes das capitais brasileiras, uma vez que onde as etnias indígenas têm forte presença, também os indigenismos registram-se em maior número, além do fato de as fronteiras geográficas também influenciarem o falar dos habitantes dessas cidades, como é o caso de São Luís que, apesar de pertencer à região Nordeste, evidenciou a alta produtividade de indigenismos tal qual as capitais da região Norte (18 registros nessa capital). Segue, pois, a Carta Linguística I:



Realizadas essas considerações, a continuidade do trabalho dá-se com a análise das unidades lexicais de base indígena relacionadas às três áreas semânticas selecionadas por este estudo, respeitando-se, para tanto, a ordem das questões do Questionário Semântico-Lexical do Atlas Linguístico do Brasil (QSL/ALiB). Antes, porém, esta parte que se dedicou às questões gerais de registro de indigenismos atém-se, finalmente, a considerações sobre o *corpus* analisado, no que diz respeito às respostas válidas e não válidas.

Isso porque foram analisadas, na continuidade do trabalho, apenas as unidades lexicais consideradas válidas como respostas para as perguntas do QSL/ALiB em questão. Uma unidade léxica foi considerada válida quando designar o referente correspondente à questão em pauta. Nesse sentido, são tidos como válidos apenas itens lexicais que, segundo sua dicionarização e/ou a explicação fornecida pelos informantes, correspondam aos semas contemplados pelos textos das perguntas do QSL/ALiB.

4.1 ANÁLISE DE DADOS DA ÁREA SEMÂNTICA ATIVIDADES AGROPASTORIS

Dentre as 25 perguntas dessa área semântica, as 21 unidades léxicas de base indígena foram auferidas como respostas para 12 delas, a saber:

| QSL | Formulação da pergunta | Indigenismo |
|------------|--|--------------------|
| 39 | “As frutas menores que a laranja, que se descascam com a mão, e, normalmente, deixam um cheiro na mão? Como elas são?” | <i>Carioquinha</i> |
| 40 | “O grão coberto por uma casquinha dura, que se come assado, cozido, torrado ou moído?” | <i>Amendoim</i> |
| 41 | “Umaz florezinhas brancas com miolo amarelinho, | <i>Maracujá</i> |

| | | |
|----|--|---|
| | ou florezinhas secas que se compram na farmácia ou no supermercado e servem para fazer um chá amarelinho, cheiroso, bom para dor de barriga de nenê/bebê e até de adulto e também para acalmar?” | |
| 43 | “Duas bananas que nascem grudadas?” | <i>Bananas cuins</i> |
| 44 | “A ponta roxa do cacho da banana?” | <i>Mangará</i> <i>Mangarataia</i> <i>Maracá</i> |
| 45 | “Quando se vai colher o milho, o que é que se tira do pé?” | <i>Soca/soca de milho</i> |
| 46 | “Quando se tira da ____ (cf. item 45) todos os grãos do milho, o que sobra?” | <i>Tamboeira</i> |
| 47 | “Depois que se corta o pé de arroz ou de fumo, ainda fica uma pequena parte no chão. Como se chama essa parte?” | <i>Soca</i> <i>Soqueira</i> |
| 50 | “Aquele raiz branca por dentro, coberta por uma casca marrom, que se cozinha para comer?” | <i>Mandioca</i> <i>Aipim</i> <i>Macaxeira</i> |
| 51 | “Uma raiz parecida com ____ (cf. item 50) que não serve para comer e se rala para fazer farinha (polvilho, goma)?” | <i>Mandioca</i> <i>Tapioca</i> <i>Jerimum</i> <i>Maniçoba</i> <i>Maniva</i> <i>Aipim</i> |
| 57 | “Aqueles objetos de vime, de taquara, de cipós trançados, para levar batatas (mandioca, macaxeira, aipim, etc.), no lombo do cavalo ou do burro?” | <i>Jacá</i> <i>Uru</i> <i>Jamaxim</i> |
| 58 | “E quando se usam objetos de couro, com tampa, para levar farinha, no lombo do cavalo ou do burro?” | <i>Uru</i> <i>Jamaxim</i> <i>Jacá</i> |

Quadro IV: Indigenismos documentados na área semântica *atividades agropastoris* nas capitais brasileiras

Dentro do universo de 308 itens lexicais auferidos como respostas para as 25 perguntas, os indigenismos correspondem a 6,8% das ocorrências. A representatividade das línguas de origem identificadas no conjunto do *corpus*

relativo à área semântica em exame pode ser observada no Gráfico III, na sequência:

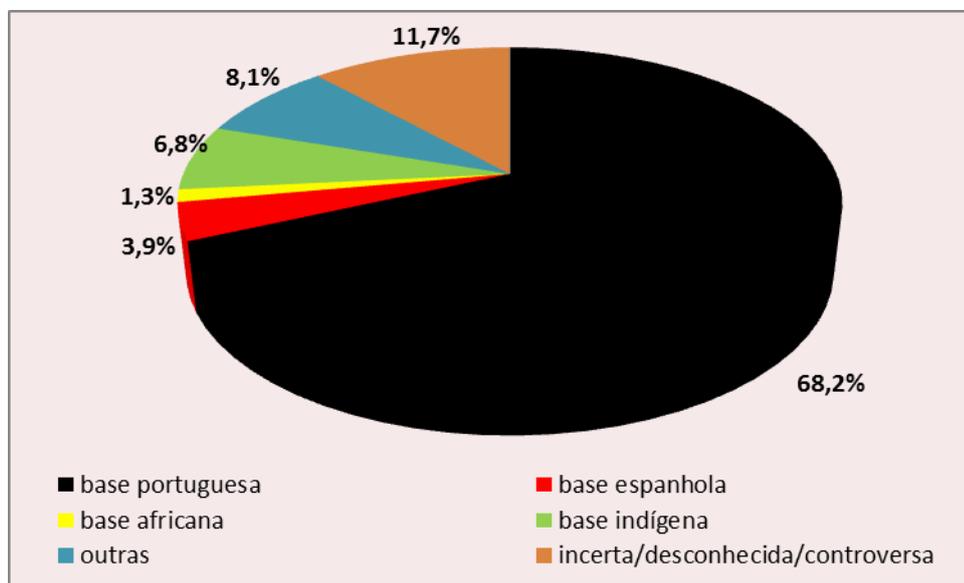


Gráfico III – Distribuição das unidades lexicais documentadas na área semântica das *atividades agropastoris* de acordo com a base linguística

Nota-se que o total de 21 indigenismos representa, no universo das perguntas relativas às *atividades agropastoris*, percentual menor que no cômputo geral das três áreas semânticas (8,4%), e que a língua portuguesa foi mais marcante nessa área (68,2% entre os dados relativos às *atividades agropastoris* contra 61,1% no total geral de respostas). Isso se deve ao fato de, na próxima área semântica a ser analisada, a *fauna*, ter sido registrado o maior percentual de indigenismos no universo pesquisado (9,3%).

Todavia, na análise dos dados foram descartadas, como respostas não válidas, as seguintes unidades léxicas: *jerimum*, *tapioca*, *maracujá* e *jamaxim* que, apesar de serem todas de origem tupi, segundo as obras pesquisadas, designam referentes diversos dos requisitados pelas perguntas do QSL/ALiB. No Quadro V,

a seguir, pode-se visualizar o conceito que expressam de fato e o requisitado pela pergunta em pauta:

| Item lexical | Conceito que expressa | Conceito requisitado |
|-----------------|--|---|
| Jamaxim | “ Cesto ³⁸ longo de três lados, ger. De trançado hexagonal, que alguns indígenas brasileiros (p.ex., os guajajara) carregam nos ombros ou prendem na testa por uma alça, e que repousa sobre suas costas” (HOUAISS, 2001). | Objetos de vime, de taquara, de cipó trançados, para levar batatas (mandioca, macaxeira, aipim, etc.), no lombo do cavalo ou do burro. |
| Jerimum | “Abóbora” (FERREIRA, 2004). | Raiz parecida com a mandioca que não serve para comer e se rala para fazer farinha (polvilho, goma). |
| Maracujá | “Design. Comum a diversas plantas do gên. <i>Passiflora</i> , da fam. Das passifloráceas, trepadeiras de frutos comestíveis e com propriedades calmantes ” (HOUAISS, 2001). | Florezinhas brancas com miolo amarelinho, ou florezinhas secas que compram na farmácia ou no supermercado e servem para fazer um chá amarelinho, cheiroso, bom para dor de barriga de nenê/bebê e até de adulto e também para acalmar. |
| Tapioca | “ Fécula comestível, extraída das raízes da mandioca ou do aipim, com que se preparam pratos doces e salgados ” (HOUAISS, 2001). | Raiz parecida com a mandioca que não serve para comer e se rala para fazer farinha (polvilho, goma). |

Quadro V: Respostas não válidas em relação ao referente requerido pela pergunta na área semântica *atividades agropastoris*

Fonte: Banco de dados ALiB

³⁸ Negritaram-se os semas partilhados pelos conceitos ou mesmo que podem ter levado à remissão, por parte dos informantes, a esses elementos diversos dos requisitados pelas perguntas da área semântica *atividades agropastoris* (QSL/ALiB).

Comparando-se as definições dicionarizadas e os conceitos expressos pelas perguntas, vê-se que *jerimum* não tem relação direta com a mandioca. Porém, as demais respostas aproximam-se de alguns semas das questões, o que pode ter motivado seus registros.

É o caso, por exemplo, de *jamaxim*. Por tratar-se de um cesto para carregar objetos, cargas, e também por ser produzido a partir de palha ou cipó, pode ter sido mencionado por 06 informantes, de Rio Branco (AC) e Boa Vista (RR). Contudo, o que o diferencia do objeto também conhecido como cesto é em quem se utiliza: o *jamaxim* é carregado por pessoas; já o cesto é colocado no lombo de animais, características essas que foram também mencionadas pelos informantes. Associação semântica semelhante pode também ter acontecido com o *maracujá*, conhecido por suas propriedades calmantes, propriedades essas partilhadas com a camomila, objeto da questão 41 do QSL/ALiB.

Já *tapioca*, registrada como resposta para a pergunta 51, tem em comum com a mandioca a sua origem, haja vista ser um produto dessa raiz, e deve ter sido associada ao sema “fazer farinha, polvilho, goma”, não pelo fato de fazer goma, mas de ser feita a partir dessa goma.

Assim, dos 21 indigenismos documentados na área semântica *atividades agropastoris*, restaram, como respostas válidas a serem analisadas, 16 unidades léxicas: *aipim*, *amendoim*, *bananas cuins*, *carioquinha*, *jacá*, *macaxeira*, *mandioca*, *mangará*, *mangarataia*, *maniçoba*, *maniva*, *maracá*, *soca*, *soqueira*, *tamboeira* e *uru*. Dada a quantidade de itens lexicais a serem analisados, foram organizadas segundo as perguntas que as geraram, seguindo, para tanto, a mesma ordem do Questionário Semântico-Lexical do Projeto Atlas Linguístico do Brasil.

Dessa maneira, a primeira unidade lexical a ser analisada é *carioquinha*, utilizada para designar “a fruta menor que a laranja, que se descasca com a mão, e, normalmente, deixa um cheiro na mão” (QSL/ALiB 39). Foi mencionada por 02 informantes paulistas para designar a fruta comumente conhecida como *mexerica*. *Carioca* advém da palavra tupi **kari’oka**, que significa casa do branco (CUNHA, 2010), e diz respeito aos indivíduos nascidos no Estado do Rio de Janeiro (FERREIRA, 2004). Em nossas investigações, não foi encontrada ligação entre *carioquinha* e *mexerica*; porém, ambos os informantes descreveram as características da fruta em questão³⁹: para o idoso de nível universitário paulista, “a *carioquinha* é pequena”, enquanto para a idosa de baixa escolaridade, a fruta com essa designação “é muito conhecida que é ruim de descascar”. Por esse motivo, apesar de nomear outro referente, a resposta consciente dos informantes tornou *carioquinha* uma resposta válida, passível de análise.

Já como designação para “o grão coberto por uma casquinha dura, que se come assado, cozido, torrado ou moído” (QSL/ALiB 40), foram registrados os itens lexicais *amendoim*, *minduí*, *mandubim*, *mundubim* e *mundubi*, aqui reunidos a *amendoim* por se caracterizarem como variantes mórficas ou fonéticas. Algumas dessas unidades lexicais estão dicionarizadas, como *mundubim/mundubi*, que figura na obra de Houaiss (2001) com remissão a *amendoim*, assim como em *mandubi*. Também Cunha (2010) registra, no verbete *amendoim*, a forma *minduim*, atribuindo-lhe a origem tupi, cujo étimo, **manu’ui**, remete ao nome de diversas plantas da família das leguminosas (CUNHA, 2010). Ferreira (2004), por sua vez, estende o sentido de *amendoim*, fornecendo-lhe a acepção de “fruto utilizado na alimentação humana, torrado ou coberto de

³⁹ Para a pergunta 39 do QSL/ALiB, há o complemento de se pedir para descrever a fruta, de modo a esclarecer as diferenças entre as várias possíveis designações.

açúcar”, mesmo conceito da questão em pauta. Altamente produtivas no âmbito dos dados analisados, *amendoim* e suas variantes foram as únicas respostas documentadas como o nome desse referente, configurando-se, pois, como uma forma categórica na língua.

Minduí, por sua vez, foi proferida apenas pelo jovem campo-grandense de baixa escolaridade. Também um jovem, porém de alta escolaridade, do Amapá, citou *mandubim*. A unidade lexical *mundubi*, por seu turno, foi mencionada pelo paraense idoso de baixa escolaridade. Como foram concebidas como variantes da forma padrão *amendoim*, as demais foram a ela agrupadas, para fins de análise diatópica e léxico-semântica dos dados.

Outra unidade lexical de origem tupi analisada foi *cuim*, utilizada para nomear as “duas bananas que nascem grudadas” (QSL/ALiB 43) (bananas *cuins* foi o registro). Mencionada por apenas uma informante, a idosa de alta escolaridade do Maranhão, *cuim*, na verdade, está dicionarizada como designação para o “ouriço-cacheiro” (FERREIRA, 2004; HOUAISS, 2001). Cunha (2010) atribui-lhe o étimo tupi **kui’ĩ**, que nomeia o “roedor da família dos eretizontídeos, espécie de cuandu”. Todavia, essa unidade lexical foi considerada válida por ter sido mencionada como primeira resposta para a questão em pauta por parte dessa informante, que depois respondeu gêmeas, demonstrando o conhecimento do referente e a certeza quanto à designação que mais utiliza. Além disso, o registro de *cuim* pode estar associado a *inconho*, advindo do tupi **y-côi** (MACHADO, 1987) e que significa “que nasce acoplado a outro (diz-se do fruto)” (HOUAISS, 2001).

Já a pergunta 44 do Questionário Semântico-Lexical do Projeto ALiB motivou o registro dos tupinismos *mangará*, *mangarataia* e *maracá*. O item lexical

mangará que nomeia “a ponta roxa do cacho da banana” é motivo de controvérsias em se tratando de etimologia, segundo as obras pesquisadas: em Machado (1987), registra-se “**ma** por **ybá**, fruto, **cará**, casculo, ou **carâ**, redondo”; Cunha (2010), por seu turno, atribui-lhe o étimo tupi **mana’ra**, que designa a “planta da família das arácias” (CUNHA, 2010). Ferreira (2004) fornece-lhe a acepção de “ponta terminal da inflorescência da bananeira” e Houaiss (2001) acrescenta-lhe as características de cores: “roxa ou castanho-avermelhada”.

O item lexical *mangarataia*, por sua vez, deriva do étimo **mangará + taia** ou **taiá**, “picante, que arde” (MACHADO, 1987) e está dicionarizado como o “açafreão da terra” (HOUAISS, 2001). Contudo, foi aqui considerada como resposta válida, dada a convicção da jovem de alta escolaridade que a mencionou em Boa Vista⁴⁰. *Maracá* também foi produtiva nessa capital, mencionada pela idosa de alta escolaridade, que ainda confirmou sua resposta: “o maracá da banana... é maracá mesmo”, apesar de estar dicionarizada como “instrumento chocalhante que era usado pelos índios nas solenidades religiosas e guerreiras” (FERREIRA, 2004). Essa unidade lexical também tem sido interpretada como de dois étimos distintos, ambos, porém, advindos do tupi: **ma’râ** “fingimento ou barulho” + **acâ** “cabeça” (MACHADO, 1987) e **mara’ka** (CUNHA, 2010).

Deve-se considerar ainda a hipótese de *maracá* poder ser uma variante fonética de *mangarataia*. Reforça essa hipótese o fato de *maracá* ter sido registrada na mesma localidade que *mangarataia*.

Soca, a próxima unidade lexical a ser analisada, foi registrada em duas das 12 perguntas que geraram indigenismos: na 45, que busca designativos para o referente também conhecido como espiga de milho, e na 47, centrada no

⁴⁰ Neste caso, para confirmação, a entrevistadora perguntou por mais de uma vez para saber a resposta, sendo respondido, em ambas as vezes, *mangarataia*.

seguinte conceito: “depois que se corta o pé de arroz ou de fumo, ainda fica uma pequena parte no chão. Como se chama essa parte?”. Essa unidade lexical, segundo Cunha (2010), advém do étimo tupi **‘soka**, e diz respeito à “segunda produção de cana-de-açúcar”. Em Houaiss (2001) e Ferreira (2004), ela é registrada como regionalismo/brasileirismo do Nordeste referindo-se à segunda colheita do fumo e, no Espírito Santo, à segunda colheita de arroz. Vê-se que essas acepções dicionarizadas relacionam-se diretamente ao referente de que se busca o nome na pergunta 47, mas não pela questão 45, haja vista que em nenhuma das obras pesquisadas *soca* refere-se à espiga de milho. Todavia, é importante registrar a alta produtividade de *soca* como resposta para espiga de milho na cidade de Florianópolis, capital de Santa Catarina. Nessa localidade, 07 dos 08 informantes a mencionaram (à exceção da informante idosa de baixa escolaridade), demonstrando que, nessa capital, *soca* é o mesmo que espiga, o que a torna resposta válida para ambas as perguntas.

Ainda na pergunta 47 (QSL/ALiB), registrou-se também *soqueira*, aqui associada à *soca*. Segundo Houaiss (2001), *soqueira* deriva de *soca* + o sufixo – eira.

Discutidas as designações vinculadas à pergunta 45 e, por associação de respostas, a 47, na sequência, analisa-se *tamboeira*, unidade lexical registrada como resposta para a pergunta 46 do Questionário ALiB (Semântico-Lexical): “quando se tira da ____ (cf. item 45) todos os grãos do milho, o que sobra?”. Mencionada por 03 dos quatro maranhenses idosos entrevistados (exceto o de baixa escolaridade), *tamboeira* advém do étimo **teomboéira** e significa “os restos; [...] os resíduos inúteis” (SAMPAIO, 1928). Por extensão de sentido, o resto, o que sobra da espiga do milho após a retirada dos grãos pode também ter

recebido esse nome. É o que comprova o registro da unidade léxica no dicionário contemporâneo de língua portuguesa, Ferreira (2004), que remete o consulente a *batuera*, brasileirismo usado para nomear o “sabugo de milho”.

Com relação à pergunta 50 do QSL/ALiB, que busca nomes para “aquela raiz branca por dentro, coberta por uma casca marrom, que se cozinha para comer”, registraram-se os tupinismos *mandioca*, *macaxeira* e *aipim*. A primeira delas, inclusive, também foi utilizada para nomear o conceito expresso pela questão seguinte, a 51, que investiga a designação para “uma raiz parecida com ____ (cf. item 50) que não serve para comer e se rala para fazer farinha (polvilho, goma)”. Sobre o item lexical *mandioca*, Cunha (2010) atribui-lhe o étimo tupi **mani’oka** e acrescenta tratar-se do vocábulo tupi mais amplamente documentado em língua portuguesa. Machado (1987) insere a informação da separação do nome em “*many* = *manyba*, maniva; *oca*, talvez *oca*, *oga*, que procede, que se tira”. Segundo Houaiss (2001), trata-se da

[...] Raiz dessa planta (arbusto (*Manihot esculenta*) da família das euforbiáceas, nativo da América do Sul, de folhas membranáceas, inflorescências ramificadas e frutos capsulares, cultivado pelas raízes tuberosas, muito semelhantes às do aipim e também ricas em amido e de largo emprego na alimentação, embora sejam geralmente mais venenosas e frequentemente usadas apenas para a produção de farinha de mandioca, farinha-d’água e ração animal).

Já Ferreira (2004), além da definição, destaca a diferença entre as espécies venenosas e as comestíveis: “planta leitosa, da família das euforbiáceas (*Manihot utilissima*), cujos grossos tubérculos radiculares, ricos em amido, são de largo emprego na alimentação, e da qual há espécies venenosas, que servem para fazer farinha de mesa”. Assim, nota-se que Ferreira (2004) menciona que há ambas as espécies, as venenosas, que fazem farinha, goma, e as comestíveis, que se cozinham para comer.

Macaxeira, a outra unidade léxica utilizada para nomear a mesma raiz, advém do tupi **maka'šera**, “diabo, entre os índios do Brasil” (CUNHA, 2010) e tem, em seus registros dicionarizados, apenas a remissão aos sinônimos talvez mais propagados, mandioca e aipim (HOUAISS, 2001; CUNHA, 2010). Houaiss (2001), inclusive, acrescenta-lhe a marca de uso no Norte e Nordeste do Brasil, informação partilhada com Ferreira (2004). Essa marca de uso foi em grande parte confirmada pelos dados aqui analisados, como a análise diatópica, a ser apresentada posteriormente, demonstrou.

Também *aipim*, de étimo **ai'pĩ**, que designa “a raiz enxuta, a mandioca mansa” (SAMPAIO, 1928), em Ferreira (2004), registra-se apenas como sinônima de mandioca. Já em Houaiss (2001), constam os sinônimos “macaxeira, macaxera, mandioca, mandioca-doce, mandioca-mansa”, além da informação: “nativo do Brasil, é muito semelhante à mandioca (*Manihot esculenta*) e também cultivado, com inúmeras variedades, pelas raízes tuberosas, de elevado teor alimentício e geralmente menos venenosas.” Assim, em Houaiss (2001), comparando-se as definições de *mandioca* e *aipim*, depreende-se que esse dicionarista associa *mandioca* à raiz buscada pela questão 51 do QSL/ALiB, enquanto *aipim* seria resposta para a pergunta 50.

Reforça essa dicionarização documentada em Houaiss (2001) o registro de variantes de mandioca fornecidas como resposta para a pergunta que busca o nome da raiz parecida com a mandioca/aipim/macaxeira que se usa para fazer farinha, goma. Os informantes mencionaram, além de *aipim*, *macaxeira* e *mandioca*, *mandioca brava*, *mandioca cacau* e *mandioca paraguaia*, talvez para diferenciá-la da mandioca comestível.

E ainda para nomear essa mesma raiz foram registradas as unidades lexicais *maniçoba* e *maniva*. *Maniçoba*, segundo Cunha (2001), provém do tupi **mani'soua** (mani = mandioca; soua =folha) e designa a folha de mandioca. Já Houaiss (2001) a define como “arbusto (*Manihot caricaefolia*) da mesma família, nativo do Brasil, de caule lenhoso, folhas liras e flores em racemos ferrugíneos e tomentosos; mandioca-brava”, aproximando *maniçoba* da mandioca de que se faz farinha.

Maniva, por sua vez, mencionada por dois macapaenses, advém do étimo também tupi **manyb** (MACHADO, 1987) e designa, segundo esse autor, a “árvore de fruto (por excelência, por ser base da alimentação)”, base da alimentação essa dos indígenas que foi passada de geração em geração entre os brasileiros. Em Houaiss (2001) e Ferreira (2004), registra-se *mandioca* na entrada *maniva*, sendo que o primeiro dicionarista ainda atribui ao item léxico a marca de uso no Norte e Nordeste do Brasil, registro parcialmente confirmado pelos dados armazenados pelo Projeto Atlas Linguístico do Brasil.

E, concluindo-se a análise léxico-semântica das unidades lexicais utilizadas para nomear os referentes ligados às *atividades agropastoris* selecionadas para este estudo, estão os nomes *jacá* e *uru*, ambas registradas como designações para os referentes requisitados pelas perguntas 57 e 58 do QSL/ALiB, que documentam, respectivamente, os designativos para “aqueles objetos de vime, de taquara, de cipós trançados, para levar batatas (mandioca, macaxeira, aipim, etc.), no lombo do cavalo ou do burro” (QSL/ALiB 57) e para os “objetos de couro, com tampa, para levar farinha, no lombo do cavalo ou do burro” (QSL/ALiB 58).

De acordo com Sampaio (1928), *jacá* origina-se do vocábulo tupi **ayacá** e nomeia o “cesto, tecido de taquaras e de forma cylindrica (sic)”. Ferreira (2004) acrescenta que se trata de “espécie de cesto feito de taquara ou de cipó, e de forma variável, para conduzir carga, em geral de comestíveis, às costas de animais”. Dessa maneira, ele diz respeito ao cesto requisitado pela pergunta 57 e não ao objeto de couro com tampa, apesar de seu registro como resposta para a questão 58, quando o informante jovem de baixa escolaridade, ao ver a gravura do objeto também conhecido como *bruaca*, respondeu *jacá*.

Isso também ocorreu com *uru*, de étimo tupi **u’ru** (CUNHA, 2010), que designa o “cesto indígena com alça, feito de palha de carnaúba” (HOUAISS, 2001), registrado como resposta para ambas as questões (57 e 58 do QSL/ALiB). Apesar de dicionarizado como o cesto de palha, o informante idoso de alta escolaridade paraibano mencionou esse item lexical para nomear a bolsa de couro, repetindo duas vezes a resposta e demonstrando certeza quanto à designação⁴¹.

Assim, concluída a análise léxico-semântica das unidades léxicas registradas na área semântica das *atividades agropastoris*, nota-se, dentre outras conclusões, que a menção de certas unidades lexicais as aproxima das acepções dicionarizadas na maior parte dos casos, mas também se verificou que há inovações linguísticas quanto à designação de certos referentes, como é o caso do uso de *soca* para nomear o sabugo do milho, por exemplo, revelando que a língua comporta-se mesmo como um organismo vivo, em constante mudança e adequação, motivadas por contingentes sociais, culturais, histórico-geográficos, dentre outros.

⁴¹ O informante, inclusive, deu como respostas para a questão 57 as unidades lexicais *cesto* e *caçuá*, demonstrando que nomeia de forma diferente ambos os referentes, além de mostrar ainda, segundo nosso entendimento, que difere ambos os objetos e os conhece.

E, findada essa primeira parte da análise, a continuidade do trabalho se dá agora com o estudo diatópico das unidades lexicais consideradas como respostas válidas para esta área semântica. Os 16 tupinismos foram registrados nas cinco regiões político-administrativas do Brasil; todavia, a documentação a distribuição dessas unidades lexicais nas capitais registrou peculiaridades, tal qual se pode visualizar por meio do Quadro VI a seguir:

| QSL/pergunta | Indigenismo | N | NE | CO | SE | S |
|--|-------------------------------|---|----|----|----|---|
| 39. “As frutas menores que a laranja, que se descascam com a mão, e, normalmente, deixam um cheiro na mão? Como elas são?” | <i>Carioquinha</i> | | | | X | |
| 40. “O grão coberto por uma casquinha dura, que se come assado, cozido, torrado ou moído” | <i>Amendoim</i> | X | X | X | X | X |
| 43. “Duas bananas que nascem grudadas” | <i>Bananas cuins</i> | | X | | | |
| 44. “A ponta roxa do cacho da banana” | <i>Mangará</i> | X | X | | | |
| | <i>Mangarataia</i> | X | | | | |
| | <i>Maracá</i> | X | | | | |
| 45. “Quando se vai colher o milho, o que é que se tira do pé?” | <i>Soca/soca de milho</i> | | | | | X |
| 46. “Quando se tira da ____ (cf. item 45) todos os grãos do milho, o que sobra?” | <i>Tamboeira</i> | | X | | | |
| 47. “Depois que se corta o pé de arroz ou de fumo, ainda fica uma pequena parte no chão. Como se chama essa parte?” | <i>Soca/soqueira</i> | X | | | X | |
| 50. “Aquele raiz branca por dentro, coberta por uma casca marrom, que se cozinha para comer” | <i>Mandioca</i> | X | X | X | X | X |
| | <i>Aipim</i> | X | X | X | X | X |
| | <i>Macaxeira</i> | X | X | X | X | X |
| 51. “Uma raiz parecida com ____ (cf. item 50) que não serve para comer e se rala para fazer farinha (polvilho, goma)” | <i>Mandioca</i> | X | X | X | X | X |
| | <i>Maniçoba</i> | X | | | | |
| | <i>Maniva</i> | X | | | | |
| | <i>Aipim</i> | | X | X | X | |
| | <i>Macaxeira</i> | X | X | X | | |
| 57. “Aqueles objetos de vime, de taquara, de cipós trançados, para levar batatas (mandioca, macaxeira, aipim, etc.), no lombo do cavalo ou do burro” | <i>Jacá</i> | X | X | X | X | X |
| | <i>Uru</i> | | X | | | |

| | | | | | | |
|---|------|---|---|--|--|--|
| 58. "Objetos de couro, com tampa, para levar farinha, no lombo do cavalo ou do burro" | Jacá | X | | | | |
| | Uru | | X | | | |

Quadro VI – Distribuição dos indigenismos na área semântica das *atividades agropastoris* documentados nas capitais brasileiras, de acordo com a região do Brasil

Nota-se, a partir dos dados do Quadro VI, que as regiões portam-se de maneira distinta no tocante aos indigenismos registrados. As regiões Norte e Nordeste, tal qual ocorre no montante geral, lideram os registros (Nordeste – 13; Norte – 14), seguidas pelas regiões Sudeste (09 ocorrências), Sul (07 indigenismos documentados) e Centro-Oeste (08 ocorrências). Esses dados corroboram as informações apresentadas no quadro geral das respostas⁴², que traziam o histórico indígena mais presente nas regiões de maior documentação de indigenismos e demonstrava que a baixa produtividade no centro-sul do Brasil poderia se dever a fatores de imigração, tanto o mais recente povoamento da região Centro-Oeste quanto à presença de etnias diversas das do Sudeste e do Sul.

Além disso, pode-se notar também que determinadas unidades lexicais disseminaram-se por todas as regiões brasileiras, como *amendoim*, *mandioca*, *macaxeira*, *aipim* e *jacá*, assim como certas unidades lexicais concentram-se em áreas específicas, como *uru* e *soca*, por exemplo. Já quanto aos itens léxicos e sua produtividade, também há resultados bem distintos, havendo unidades altamente recorrentes e outras até com registros únicos. A organização dos itens lexicais analisados por ordem de produtividade evidencia o seguinte panorama:

| Item lexical | Número de ocorrências (base de dados/ALiB) |
|------------------------|--|
| <i>Mandioca</i> | 256 |

⁴² Cf. Quadro VI, páginas 123-124.

| | |
|----------------------|-----|
| Amendoim | 200 |
| Macaxeira | 128 |
| Aipim | 77 |
| Mangará | 31 |
| Jacá | 18 |
| Soca | 15 |
| Tamboeira | 3 |
| Carioquinha | 2 |
| Maniçoba | 2 |
| Maniva | 2 |
| Uru | 2 |
| Bananas cuíns | 1 |
| Mangarataia | 1 |
| Maracá | 1 |

Quadro VII – Produtividade dos indigenismos registrados na área semântica das atividades agropastoris nas capitais brasileiras

Vê-se, comparando-se os Quadros VI e VII, que as unidades lexicais mais produtivas são também as que se propagaram por todo o Brasil, à exceção de *mangará*, que teve alta produtividade apenas nas regiões Norte e Nordeste (16 no Norte e 15 no Nordeste), mas não ocorreu nas demais áreas político-administrativas brasileiras.

E dentre os itens léxicos mais recorrentes, encontram-se *mandioca*, *macaxeira* e *aipim*. Documentados como respostas para as perguntas 50 e 51 do QSL/ALiB, seus registros mais fortes foram para a questão 50 do referido Atlas (285 ocorrências somadas, contra 189 na pergunta 51). Isso porque *macaxeira* e *aipim* tiveram pouca ocorrência para a segunda pergunta em pauta (06 ocorrências para *aipim* e 15 para *macaxeira*), que concentrou *mandioca* como resposta (168 registros).

A distribuição diatópica das respostas para a pergunta 50 (QSL/ALiB), que investiga os designativos para “aquela raiz branca por dentro, coberta por uma casca marrom, que se cozinha para comer”, demonstrou que existem áreas dialetais bem demarcadas no espaço brasileiro, para além das fronteiras geográficas.

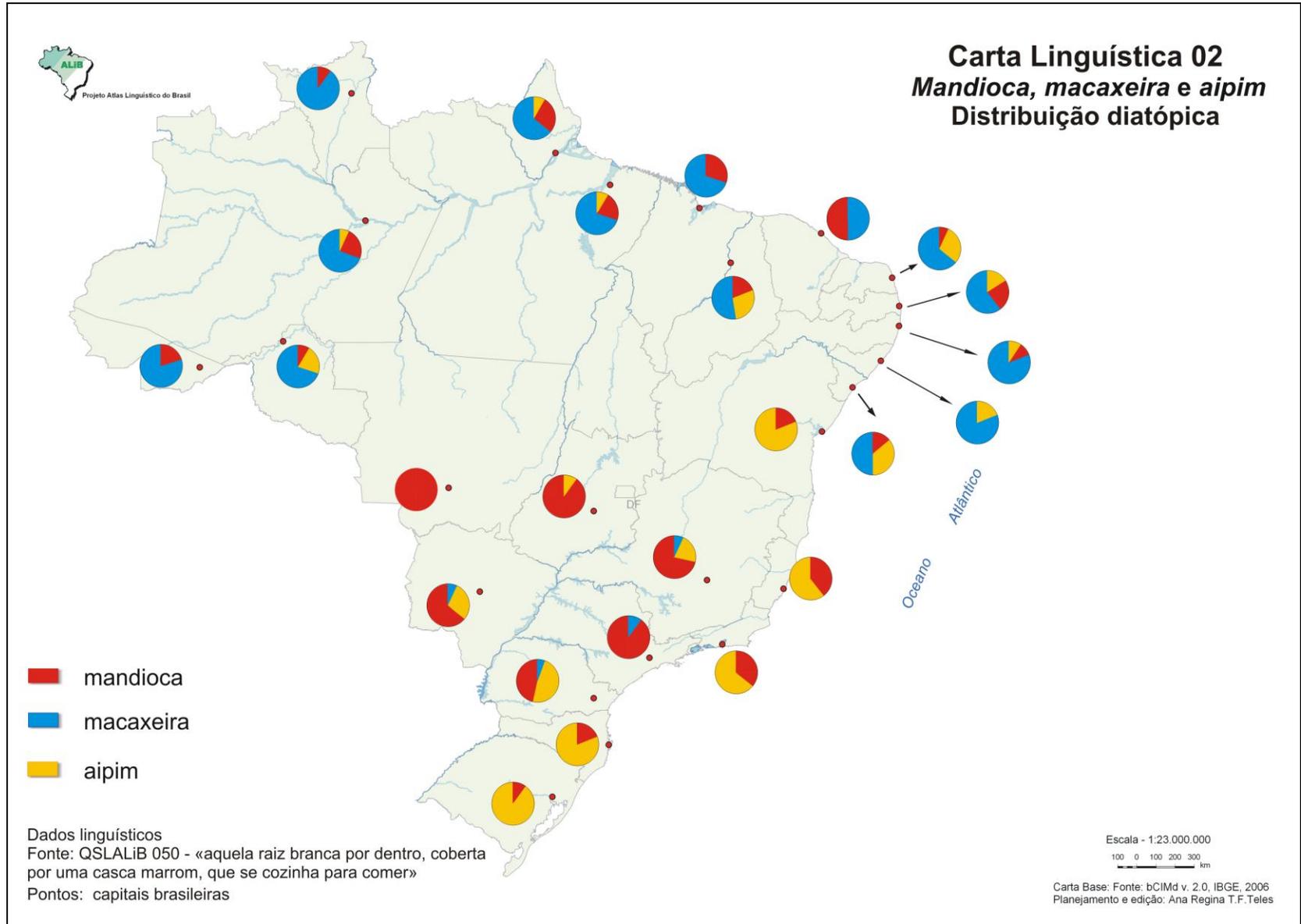
Mandioca, a unidade lexical mais produtiva, por exemplo, está em todas as regiões brasileiras, com produtividade, porém, bem distinta de acordo com a localidade que se estuda. Dominante na fala dos informantes da região Centro-Oeste (100% das ocorrências em Cuiabá, 88,9% em Goiânia e 66,7% em Campo Grande), ela sofre um quase apagamento em boa parte da região Sul (Florianópolis – 20% - e Porto Alegre – 11,1%), assim como em certas capitais nordestinas – em Natal, por exemplo, teve apenas 6,1% das ocorrências.

Já a documentação de *macaxeira*, que Houaiss (2001) e Ferreira (2004) registram com marca de uso regional no Norte e Nordeste do Brasil, confirmou essa característica, tendo em vista que, nessas regiões, foi o item lexical mais produtivo para nomear a raiz em questão (exceto em Salvador, na Bahia, em que *aipim* alçou 80% de produtividade, não tendo sido, inclusive, registrada a forma *macaxeira*). No Norte, ele alcançou os maiores percentuais de produtividade (a localidade que mais documentou *macaxeira*, no universo da questão 50, foi Boa Vista, com 88,9% dos registros). Já o Nordeste brasileiro demonstra que, apesar de ter como típico esse designativo, há outras unidades lexicais sendo incorporadas no léxico dos habitantes daquela região. Confirma essa afirmação a produtividade de *aipim* em Salvador, capital do Estado que marca a fronteira entre a região Nordeste e a Sudeste do Brasil, bem como a presença também de *aipim* em Natal, em Teresina e em Aracaju, assim como de

mandioca em São Luís, em João Pessoa e em Fortaleza (nesta capital, *mandioca* alçou 50% de produtividade).

Aipim, por seu turno, demonstrou força no Sul do país, com percentuais elevados de ocorrência em Florianópolis e Porto Alegre (80% e 88,9%, respectivamente), assim como foi a variante lexical mais produtiva nas capitais litorâneas do Sudeste brasileiro (61,5% das ocorrências em Vitória e 63,7% no Rio de Janeiro).

A partir dessas considerações, elaborou-se uma Carta Linguística demonstrando a trajetória dialetal de *mandioca*, *macaxeira* e *aipim*, registradas como resposta para a pergunta 50 do QSL/ALiB. A Carta apresenta a presença/ausência das variantes lexicais em pauta e, em sua sequência, visualiza-se a Tabela I, com os percentuais de ocorrência em cada capital:



| CAPITAL | Mandioca | Macaxeira | Aipim |
|----------------|-----------------|------------------|--------------|
| Manaus | 23,1% | 69,2% | 7,7% |
| Rio Branco | 20% | 80% | 0% |
| Porto Velho | 9,1% | 72,7% | 18,2% |
| Boa Vista | 11,1% | 88,9% | 0% |
| Macapá | 25% | 66,7% | 8,3% |
| Belém | 18,2% | 72,7% | 9,1% |
| | | | |
| Cuiabá | 100% | 0,0% | 0,0% |
| Campo Grande | 66,7% | 6,6% | 26,7% |
| Goiânia | 88,9 % | 0,0 | 11,1 % |
| | | | |
| Fortaleza | 50% | 50% | 0% |
| Natal | 6,1% | 66,6% | 27,3% |
| Teresina | 20% | 53,3% | 26,7% |
| São Luís | 27,3% | 72,7% | 0% |
| Recife | 10% | 80% | 10% |
| João Pessoa | 23,1% | 61,5% | 15,4% |
| Maceió | 0,0% | 80% | 20% |
| Aracaju | 14,3% | 50% | 35,7% |
| Salvador | 20% | 0% | 80% |
| | | | |
| Belo Horizonte | 69,2% | 7,1% | 23,7% |
| Vitória | 38,5% | 0% | 61,5% |
| Rio de Janeiro | 36,3% | 0% | 63,7% |
| São Paulo | 88,9% | 11,1% | 0% |
| | | | |

| | | | |
|---------------|-------|------|-------|
| Florianópolis | 20% | 0% | 80% |
| Porto Alegre | 11,1% | 0% | 88,9% |
| Curitiba | 46,7% | 6,6% | 46,7% |

Quadro VIII – Percentuais de ocorrência de *mandioca*, *macaxeira* e *aipim* nas capitais brasileiras

Na Carta Linguística II e no Quadro VIII, nota-se que *macaxeira* desenha um *continuum* de altas ocorrências concentrado nas regiões Norte e Nordeste do país. Nas regiões mais abaixo, no mapa, essa unidade lexical perde espaço, sofrendo um quase apagamento nas demais regiões (no Centro-Oeste, foi registrada apenas em Campo Grande, como 6,6% das ocorrências, mesmo registro de Curitiba e, no Sudeste, teve 7,1% de documentação em Belo Horizonte e 11,1% em São Paulo). Já *mandioca* domina as ocorrências no Centro-Oeste e nas capitais dos Estados vizinhos Minas Gerais e São Paulo, enquanto *aipim* domina o litoral do Sudeste e o Sul do Brasil.

Na sequência, tem-se a unidade lexical de origem tupi *mangará*. Com 31 ocorrências concentradas no Norte e Nordeste do país, *mangará* corresponde a 23,5% das ocorrências de respostas para a pergunta 44 do Questionário Semântico-Lexical do Projeto ALiB, que busca nomes para “a ponta roxa do cacho da banana”, seguida de perto por *umbigo*⁴³ - 21,2% - e empatada com as demais designações de baixa ocorrência – 23,5%). Assim, vê-se que esse item léxico foi o mais recorrente para designar o conceito em questão, mesmo estando presente apenas em duas das cinco regiões brasileiras; além disso, o universo de respostas para a QSL/ALiB 44 agrega ainda o fato de ser esta pergunta uma daquelas que, em geral, mais

⁴³ O item lexical *umbigo* não integra o *corpus* deste estudo por ser de origem portuguesa.

possui não respostas⁴⁴ (90 ocorrências), o que pode demonstrar o afastamento, por parte dos informantes entrevistados, do referente em questão, típico do meio rural.

Mangará foi também, no universo de sua documentação, mais produtiva entre os informantes idosos (83,9%), assim como teve maior ocorrência entre os entrevistados de baixa escolaridade (67,7%), demonstrando que pode estar caindo em desuso, assim como a próxima unidade léxica a ser analisada, *jacá*, seja por estarem perdendo espaço entre os mais jovens por se registrarem na fala mais de idosos, seja por terem sido preteridas frente a outros itens lexicais, por vezes pelo afastamento do referente, por vezes por interferência do processo de escolarização.

A unidade lexical *jacá* foi utilizada mormente para nomear “aqueles objetos de vime, de taquara, de cipós trançados, para levar batatas (mandioca, macaxeira, aipim, etc.), no lombo do cavalo ou do burro” (QSL/ALiB 57). Documentada como resposta para essa pergunta em 07 das 25 capitais brasileiras investigadas (Norte: Boa Vista; Nordeste: Teresina e São Luís; Centro-Oeste: Goiânia e Campo Grande; Sudeste: Vitória e Sul: Florianópolis), além de sua vertente diatópica, o que marca essa documentação é a variável diageracional, haja vista que o maior número de suas ocorrências deu-se na fala dos idosos, como se pode verificar no Gráfico IV, a seguir:

⁴⁴ Denomina-se “não resposta” o caso em que o informante não sabe nomear o conceito em pauta. Quando isso acontece, normalmente o entrevistador refaz a pergunta, reformulando-a de diferentes formas e, quando essas alternativas são improdutivas, ainda refaz a pergunta na fase das retomadas, no final do inquérito.

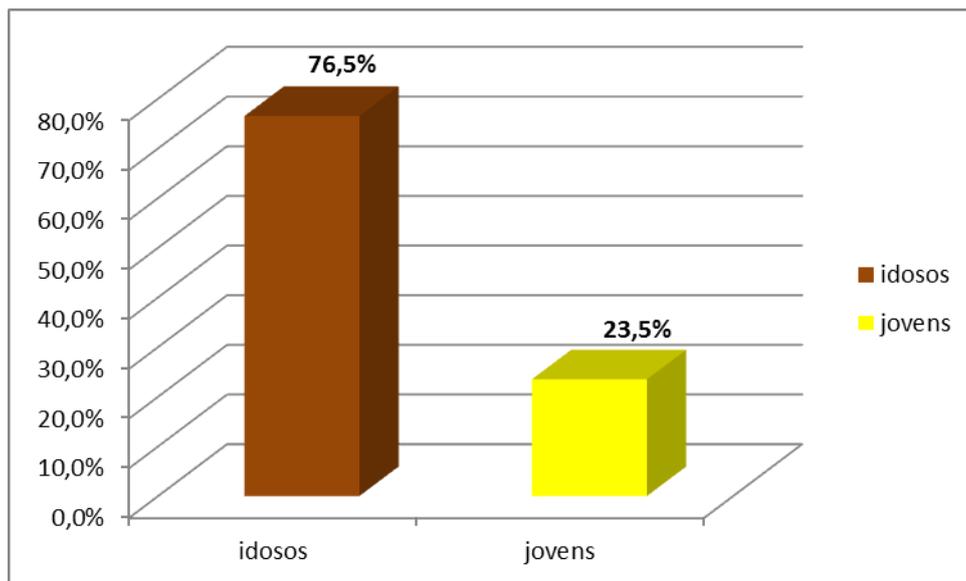


Gráfico IV – Distribuição diageracional de *jacá* nas capitais brasileiras

E, dentre os 04 jovens que preferiram *jacá* para nomear o objeto também conhecido como cesto⁴⁵, é importante ressaltar que 03 deles moram em Teresina, capital do Piauí, localidade que registrou o maior número de ocorrências dessa unidade lexical (07, que correspondem a 58,3%), não tendo sido mencionada, nessa capital, apenas pelo jovem de alta escolaridade.

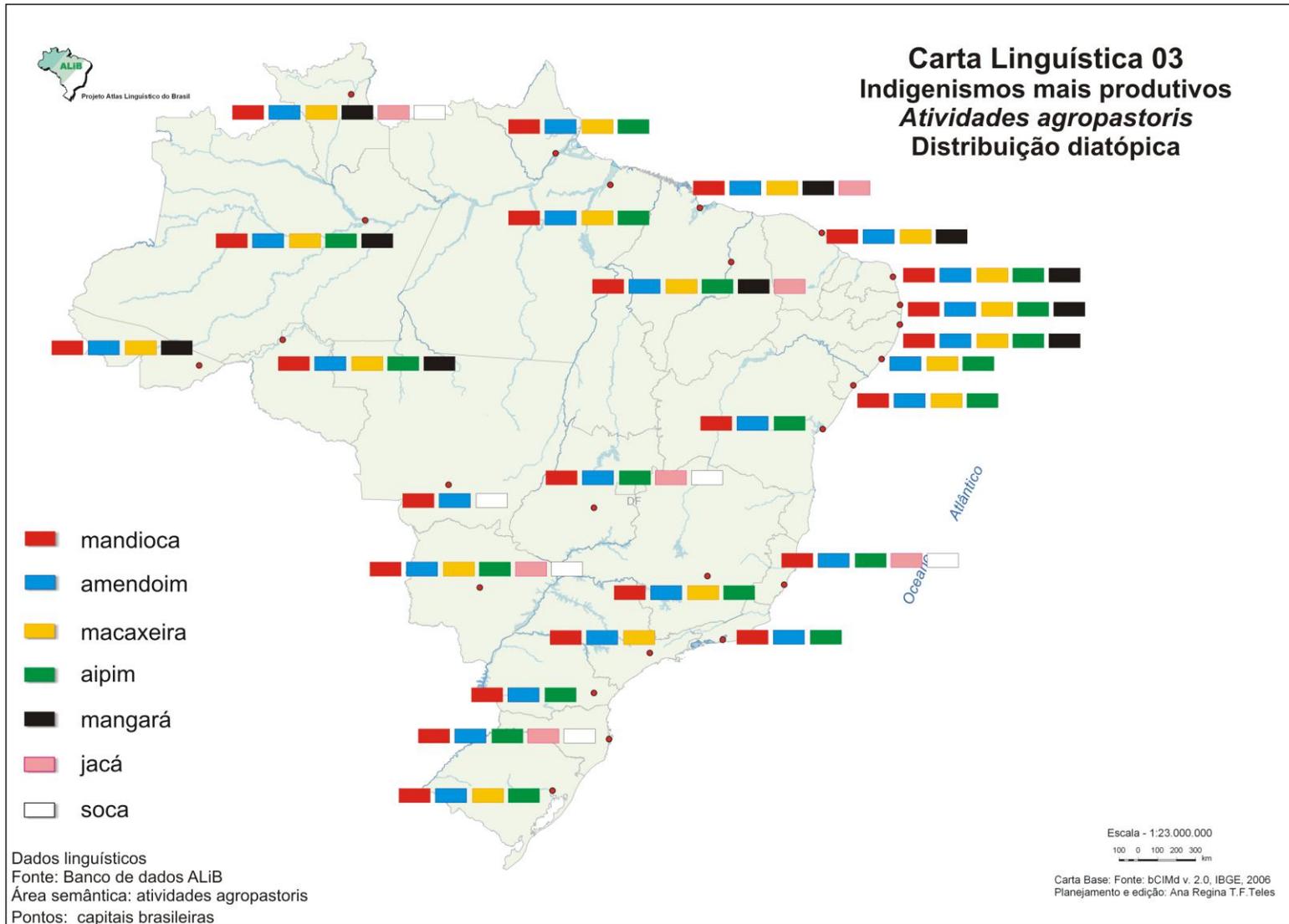
A baixa documentação de *jacá*, dentre os designativos para nomear o cesto, e sua presença mais forte entre os idosos demonstra também que esse designativo pode estar caindo em desuso, configurando-se como item lexical em vias de desaparecimento – equivale a 6,5% das ocorrências, essas lideradas por *cesto*, que aparece com 42,3% dentre o universo pesquisado. Além disso, sua maior documentação no Nordeste revela também que essa região dá mostra de certo conservadorismo em termos de norma lexical, haja vista que *jacá*, segundo Machado (1987), teve sua dicionarização já em Silva (1813). Corrobora ainda essa

⁴⁵ *Cesto* também não vigora neste trabalho por ser de base linguística portuguesa.

constatação a presença também no Norte e Nordeste de *mangará*, que, no que diz respeito à faixa etária à diatopia, assemelha-se ao caso do *jacá*.

Há que se pontuar ainda a questão de *soca*, também entre as unidades lexicais mais produtivas, como ocorre em Florianópolis, onde quase a totalidade dos informantes (07 ocorrências) mencionou *soca* como resposta para a pergunta que busca nomes para o “sabugo de milho”. Trata-se de uma marca dialetal significativa a ser considerada, além da expectativa de o futuro Atlas Linguístico do Brasil poder contribuir com dados concretos para o enriquecimento das obras lexicográficas, no que diz respeito às marcas de uso.

Concluindo-se a análise das sete unidades lexicais de origem indígena mais produtivas na área semântica das *atividades agropastoris*, faz-se um comparativo entre seu viés diatópico e as marcas de uso dicionarizadas nas obras lexicográficas. Isso porque faz parte dos objetivos específicos deste trabalho contribuir para a Lexicografia, por meio do fornecimento de informações sobre os usos de itens léxicos nas regiões investigadas pelo Projeto Atlas Linguístico do Brasil representadas pelas 25 capitais brasileiras que fazem parte de sua rede de pontos. Para isso, elaborou-se inicialmente uma carta linguística demonstrando as ocorrências de *aipim*, *amendoim*, *macaxeira*, *mandioca*, *jacá*, *mangará* e *soca* nas capitais brasileiras, para depois ser realizada a análise:



As obras aqui referidas são o dicionário eletrônico Houaiss (2001) e a versão também eletrônica do dicionário Aurélio (FERREIRA, 2004). Em ambas as obras, apenas três dessas unidades lexicais são registradas com marcas de uso: *macaxeira*, *mangará* e *soca*, como já referido anteriormente. Os dados do Projeto ALiB confirmaram a marca de uso, no Norte e no Nordeste, da unidade lexical *macaxeira*, fornecida pelas duas obras lexicográficas consultadas, pois, apesar de ter sido coletada em todas as regiões brasileiras, teve maior índice de produtividade nessas duas regiões do Brasil.

O item lexical *mangará*, por sua vez, segundo os dados aqui examinados, expandiu a sua área de uso para além da dicionarizada, haja vista que, além de ser documentada em seis das nove capitais nordestinas, confirmando, assim, o seu uso na região apontada por Houaiss (2001) e por Ferreira (2004) (regionalismo/brasileirismo do Nordeste), também foi documentada em quatro das seis capitais nortistas (Manaus, Rio Branco, Porto Velho e Boa Vista).

Já no caso da unidade lexical *soca*, registrada em Houaiss (2001) e em Ferreira (2004) como de uso regional do Nordeste, na acepção de “segunda colheita de fumo”, e do Espírito Santo como a “segunda colheita do arroz”, os dados geolinguísticos em estudo não confirmaram essas marcas no português contemporâneo, uma vez que, em Vitória (Espírito Santo), houve apenas uma ocorrência, não tendo sido documentada no Nordeste. Contudo, foi fornecida como designação do “sabugo do milho” na capital de Santa Catarina, Florianópolis, por sete dos oito falantes entrevistados, o que aponta para uma nova demarcação do espaço de uso dessa unidade lexical em acepção diversa da consagrada pelo uso na língua portuguesa.

Além dessas três unidades lexicais com marcas de uso nas obras lexicográficas pesquisadas, merece destaque a questão da distribuição espacial do item lexical *aipim*. Apesar de ter sido documentado nas cinco regiões brasileiras, esse item lexical alcançou percentuais de ocorrência consideráveis no sul do Brasil, sendo a unidade lexical mais produtiva nas três capitais sulistas (Curitiba, Florianópolis e Porto Alegre), além da capital baiana, Salvador (80%), e do Rio de Janeiro (63,7%).

Assim, comparando-se os dados lexicográficos de marcas de uso e a documentação das unidades léxicas em pauta na base de dados do Projeto Atlas Linguístico do Brasil, poderíamos traçar o seguinte panorama:

| Item lexical | Houaiss (2001) | Ferreira (2004) | Dados ALiB (capitais brasileiras) |
|------------------|---|---|--|
| Macaxeira | Regionalismo do Norte e do Nordeste do Brasil | Brasileirismo do Norte e do Nordeste | N/NE |
| Mangará | Regionalismo Nordeste | Brasileirismo Nordeste | N/NE |
| Soca | Não dicionarizado na acepção de “sabugo de milho” | Não dicionarizado na acepção de “sabugo de milho” | Sul/Florianópolis (sabugo de milho) |

Quadro IX – Marcas dialetais dicionarizadas e documentadas pelo Projeto ALiB - itens lexicais da área semântica das *atividades agropastoris*

Nota-se, pois, que os dados do projeto Atlas Linguístico do Brasil têm muito a acrescentar aos dicionários gerais da língua portuguesa quanto ao conhecimento do uso efetivo do léxico, em especial no que diz respeito às marcas de uso, os regionalismos.

Além disso, a análise léxico-semântica e diatópica aqui realizada demonstra que as pesquisas geolinguísticas estão avançando no mapeamento da

língua, tanto no verificar das acepções em uso, no revelar das escolhas lexicais dos falantes, quanto na disseminação/fixação dos itens lexicais em dado espaço e tempo.

A análise das unidades lexicais que nomeiam referentes associados às *atividades agropastoris* (QSL/ALiB) confirmou também que a língua tupi, base da Língua Geral Paulista e idioma mais propagado no Brasil Colonial, continua viva na fala dos habitantes das capitais brasileiras, tendo em vista ser essa a única base linguística indígena verificada. Segundo Melo (1981, p. 41), “das línguas indígenas do Brasil [...] o tupi [foi] aquela de maior influxo exerceu no português, porque era a mais importante, a mais falada e funcionava mesmo como espécie de ‘língua segunda’ de certos grupos aborígenes não-tupis”, o que é ratificado pelos resultados deste estudo, uma vez que se verificou que o léxico do português do Brasil foi enriquecido com um considerável contingente de dados lexicais provenientes do tupi.

Realizadas essas considerações acerca dos indigenismos relacionados à área semântica *atividades agropastoris*, o tópico seguinte é reservado à análise dos indigenismos agregados à área semântica de maior produtividade dessa categoria de item lexical, dentre as três selecionadas para este trabalho: a *fauna*.

4.2 ANÁLISE DE DADOS DA ÁREA SEMÂNTICA FAUNA

Das respostas fornecidas para as 25 questões tomadas como fonte de dados, foram auferidos 311 itens lexicais, sendo 30 de origem tupi documentados em 10 das questões. Como nas demais áreas semânticas, a origem portuguesa lidera entre as unidades lexicais documentadas (181 registros). Contudo, na área da *fauna*,

houve o maior índice de unidades lexicais de origem tupi (21 itens léxicos nas *atividades agropastoris* e 08 na *alimentação e cozinha*). Essa realidade, bem como o percentual de ocorrências de itens lexicais oriundos de outras línguas, está visualizada no Gráfico V, a seguir:

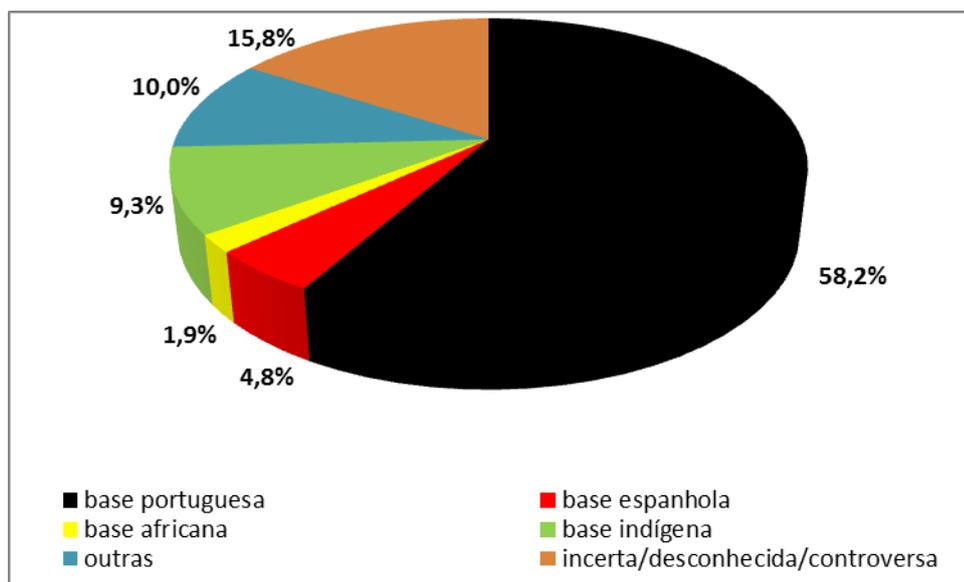


Gráfico V – Distribuição dos itens lexicais documentados na área semântica da *fauna* de acordo com a base linguística

Nota-se que as variantes lexicais de origem indígena representam 9,3% do total dos dados relativos à *fauna*, ou seja, trata-se da segunda mais produtiva do *corpus* aqui analisado. Como assinalado anteriormente, ocorreu, nessa área semântica, o maior índice de indigenismos. Isso pode ser justificado pelo fato de a fauna agregar elementos tipicamente locais, como os animais, muitos deles pertencentes a espécies distintas das já conhecidas pelos portugueses. Esses 30 itens lexicais de origem indígena são apresentados a seguir, acompanhados das respectivas perguntas que os motivaram:

| QSL | Pergunta | Indigenismo |
|-----|---|--------------|
| 64 | “A ave preta que come animal morto, podre?” | <i>Urubu</i> |

| | | |
|----|--|---|
| 68 | “A ave de penas coloridas que, quando presa, pode aprender a falar?” | <i>Maretaca</i> <i>Arara</i> <i>landaia</i> <i>Tangará</i> <i>Curica</i> <i>Maracanã</i> <i>Ajuru</i> |
| 69 | “Uma galinha sem rabo?” | <i>Nambu</i> <i>Mutum</i> <i>Socó</i> |
| 71 | “O bicho que solta um cheiro ruim quando se sente ameaçado?” | <i>Gambá</i> <i>Jaratataca/tacaca/maria tacaca</i> <i>Mucura</i> <i>Timbu</i> <i>Tamanduá</i> <i>Coati</i> <i>Saruê</i> |
| 82 | “O animal que tem uma perna mais curta e que puxa de uma perna?” | <i>Capenga</i> <i>Caruado</i> |
| 83 | “Um tipo de mosca grande, esverdeada, que faz um barulhão quando voa?” | <i>Mutuca</i> |
| 84 | “Um bichinho que se gruda nas pernas das pessoas quando elas entram num córrego ou banhado?” | <i>Mutuca</i> |
| 86 | “Aquele bichinho branco, enrugadinho, que dá em goiaba, em coco?” | <i>Tapuru</i> |
| 87 | “Aquele bicho que dá em esterco, em pau podre?” | <i>Tapuru</i> <i>Cupim</i> <i>Emboá</i> <i>Turu</i> |
| 88 | “Aquele inseto pequeno, de perninhas compridas, que canta no ouvido das pessoas, de noite?” | <i>Muriçoca</i> <i>Maruim</i> <i>Carapanã</i> |

Quadro X – Indigenismos documentados na área semântica da *fauna*, de acordo com a pergunta do Questionário Semântico-lexical/ALiB

Vê-se que duas das 10 perguntas listadas se destacaram quanto ao número de indigenismos (perguntas 68 e 71), cada uma reunindo 07 unidades

léxicas de origem indígena, que nomeiam espécies tipicamente sul-americanas, o *gambá* e o *papagaio*. A continuidade do trabalho discutirá os nomes atribuídos a esses animais. Todavia, antes de analisar as unidades lexicais obtidas como respostas para as perguntas pertencentes à área da *fauna*, a exemplo do ocorrido com a área semântica cujos dados foram analisados anteriormente, as unidades lexicais apuradas foram examinadas do ponto de vista da sua validade ou não em função do conceito nomeado. Dentre os 30 indigenismos aqui catalogados, foram consideradas respostas não válidas as unidades lexicais *tangará* (para nomear o papagaio – QSL/ALiB 68), *mutum* (QSL/ALiB 69) e *tamanduá* e *coati* (QSL/ALiB 71), além de *emboá* (QSL/ALiB 87), uma vez que, apesar de, por vezes, nomearem referente semelhante ao requerido, na maior parte dos casos, afastam-se por completo da pergunta em pauta, bem como a argumentação dos informantes que a mencionaram não sustenta a validade da resposta.

O item lexical *tangará*, por exemplo, nomeia as “aves passeriformes da família dos píprídeos, encontradas em toda a América do Sul; machos geralmente coloridos, especialmente a cabeça, e fêmeas verdes” (HOUAISS, 2001). Assim, a provável associação estabelecida pelo falante idoso cuiabano entre essa ave e a normalmente nomeada por *papagaio* pode ter sido motivada pelo colorido das penas dessas aves. Além disso, segundo o mesmo informante, o *tangará* não fala, mas gorjeia.

Para nomear o conceito expresso pela pergunta 69 (QSL/ALiB, que busca designativos para a “galinha sura”) e também considerada como resposta não válida, tem-se *mutum*. Segundo Ferreira (2004), trata-se de designação “comum a várias aves galiformes cracídeas, gênero *Crax*, de penas da crista curvas na extremidade”. Nota-se ser uma ave distinta da galinha sem rabo. E,

além da dicionarização diversa, também a ocorrência item lexical o invalida como resposta, haja vista que o idoso acreano de alta escolaridade demonstrou dúvida ao responder a pergunta – “É, chama-se mutum, né?”.

Já as respostas consideradas como não válidas para a questão 71 (que busca os designativos para o animal também conhecido como gambá) foram *tamanduá* e *coati*. Segundo Houaiss (2001), os tamanduás “têm focinho longo e tubular, dentes ausentes, língua longa e pegajosa, e grandes garras nas patas anteriores, usadas principalmente para abrir formigueiros e cupinzeiros”, o que os afasta do referente em questão. *Coati*, por seu turno, segundo Sampaio (1928, p. 308), deriva do étimo **qua-ti** e nomeia o animal “que traz riscas pelo corpo”. Já o dicionário de Cunha (1998) não contempla essa unidade lexical na sua nomenclatura. Ferreira (2004) acrescenta tratar-se de “mamífero carnívoro [...]. A coloração, em geral, é cinzento-amarelada, porém muito variável, havendo indivíduos quase pretos e outros bastante avermelhados, focinho e pés pretos, cauda com 55 cm, com sete a oito anéis pretos”.

Possivelmente, as semelhanças entre a cauda longa e a listra no dorso presentes nos dois referentes (*gambá* e *coati*) podem ter levado a informante idosa com Ensino Fundamental do Amapá a utilizar a unidade léxica em causa para nomear o “bicho que solta cheiro ruim quando se sente ameaçado”, característica essa atribuída apenas ao animal conhecido como *gambá*.

A última unidade lexical indígena catalogada como resposta na área semântica da *fauna* e considerada como resposta não válida foi *emboá/imbuá* (QSL/ALiB 87). Em Houaiss (2001), no verbete *imbuá*, há a remissão a *embuá* e em Ferreira (2004) não há registro de *imbuá*. Dessa maneira, aqui foram consideradas variantes.

Cunha (1998) não registra essa unidade lexical. Já Sampaio (1928, p. 248) atribui a ela o étimo tupi **a-mbo-ã**, que “designa a lagarta felpuda entre os índios. Outros dão este nome à centopeia”. Em Houaiss (2001) encontramos especificações da espécie: “vivem em lugares úmidos e escuros, embaixo de pedras ou troncos em decomposição, e evitam a luz”, além de se apresentarem sinônimos também compartilhados por Ferreira (2004): “bicho-bola, bicho-de-ouvido, caramuji, gongolo, gongolô, piolho-de-cobra, surrupeio”. Nota-se que *imbuá* nomeia conceito diferente do apresentado pela pergunta em causa. Uma hipótese para a menção dessa unidade lexical para designar o “bicho de esterco” pode ser a semelhança entre o *habitat* dos dois insetos, provável associação realizada pela informante idosa de baixa escolaridade, habitante do Acre, e pelo jovem paraense de baixa escolaridade, para utilizar *imbuá* para nomear o referente mais conhecido na região como *tapuru*.

Realizadas as considerações acerca das 05 unidades lexicais descartadas do *corpus* deste estudo pelas razões expostas, este trabalho ocupa-se agora do exame das demais 25 unidades léxicas de base indígena documentadas, inicialmente com o estudo léxico-semântico e depois com o diatópico. Antes, porém, é necessário mencionar que, para fins de análise léxico-semântica, as unidades lexicais foram distribuídas em 03 grupos, definidos segundo as classes taxionômicas a que pertencem os referentes nomeados, no âmbito da Zoologia: **aves** (*ajuru, arara, curica, iandaia, maracanã, maritaca, nambu, socó e urubu*); **mamíferos** (*capenga, caruado, gambá, jaratataca/tacaca/maria tacaca, mucura, saruê e timbu*) e **insetos** (*carapanã, cupim, maruim/muruim, muriçoca, mutuca/motuca, tapuru/taparu/tapuruzim e turu/toru*).

Em termos de produtividade, o tupinismo com maior número de ocorrências nas capitais do Brasil foi *urubu*, que foi fornecido como resposta para a pergunta 64 do QSL/ALiB (“ave preta que come animal morto, podre”) pelos 200 informantes entrevistados nas 25 capitais pesquisadas⁴⁶. Em Sampaio (1928, p.339), a forma *urubu* é definida como “ave voraz” e origina-se do tupi **uru’wu**. Segundo Houaiss (2001), *urubu* designa “diversas aves ciconiformes, gên. *Coragyps* e *Cathartes*, da família dos catartídeos” que “têm cabeça e pescoço nus e alimentam-se de carne em putrefação”. O item lexical *urubu* disseminou-se por todo o Brasil, como vêm atestando os dados geolinguísticos recolhidos em diferentes regiões do Brasil, a partir da segunda metade do século XX.

Ainda no grupo das aves, a unidade lexical *ajuru* foi apurada como designativa da “ave de penas coloridas que quando presa pode aprender a falar” (QSL/ALiB 68). Segundo Sampaio (1928), trata-se de uma forma derivada do étimo tupi **a-jurú** que nomeia “boca de gente, ou que tem fala como gente”, designando, posteriormente, o próprio papagaio. Houaiss (2001) acrescenta-lhe a marca de uso de regionalismo da Amazônia, marca parcialmente confirmada pelos dados do Projeto Atlas Linguístico do Brasil, uma vez que, apesar de ter apenas 01 ocorrência na base de dados, essa unidade lexical foi mencionada por um morador do Macapá, capital que se situa na região amazônica.

Arara também foi mencionada para designar a “ave de penas coloridas que quando presa pode aprender a falar” (pergunta 68 do QSL/ALiB). Trata-se de um nome tupi derivado de **a’rara**, o “nome comum a diversas aves de grande porte da família dos psitacídeos”, segundo Cunha (1998). Sampaio (1928, p.159),

⁴⁶ Os resultados deste estudo complementam o de Silva e Aguilera (2007), sobre os dados documentados como resposta para essa pergunta, em 13 capitais das regiões Norte, Centro-Oeste, Sul e Sudeste, capitais essas também contempladas por este trabalho.

por seu turno, informa tratar-se de “voz onomatopaica com que se designam os grandes papagaios”. Também em Sampaio (1928, p.156), há o registro de **ará** como designação de “papagaios grandes”. A posição de Cunha (1998) é partilhada por lexicógrafos que produziram dicionários da língua portuguesa contemporânea, como Houaiss (2001) e Ferreira (2004). O primeiro define *arara* como designação de “algumas aves psitacíformes da família dos psitacídeos (*Anodorhynchus*, *Ara* e *Cyanopsitta*), que vivem na América Latina, possuem grande porte e são dotadas de bico alto, recurvado e de cauda longa”. Já o segundo acrescenta à definição do lexicógrafo anterior o sema “se alimentam de frutas e sementes em geral”.

Aproximando essas definições do item lexical *arara* do conceito expresso pela pergunta 68 do QSL/ALiB, como também as informações fornecidas por Sampaio (1928), têm-se os depoimentos dos informantes que mencionaram essa unidade lexical como designativa da ave em questão e que argumentam que essa ave também é colorida e aprende a falar, tal qual o papagaio. De acordo com o jovem goiano de alta escolaridade, por exemplo, a *arara* aprende a falar e ele conhece, inclusive, uma ave dessas que fala – “INQ. – A arara aprende a falar será? INF.- Aprende. INQ⁴⁷. – Arara? Ah, é? Não sabia não. INF. – Sabia não? O vizinho lá de casa tinha”. Também a idosa de Porto Velho de alta escolaridade confirma que a *arara* também pode falar, apesar de menos que o papagaio.

Dessa maneira, apesar de não ser designativo para o papagaio propriamente dito, também *arara* nomeia uma ave de penas coloridas que aprende a falar, se em contato direto com o ser humano. E, assim sendo, além da

⁴⁷ INQ é a abreviatura de inquiridor e identifica o pesquisador do ALiB responsável pela aplicação do questionário. Já INF é a sigla referente ao informante entrevistado.

justificativa da resposta de diversos informantes, também a alta produtividade desse item lexical (18 ocorrências), no *corpus* pesquisado, caracteriza-o como resposta válida.

Nomeando também o conceito expresso pela pergunta 68 do QSL/ALiB registrou-se a unidade lexical *curica* que, em Sampaio (1928, p. 228), designa a ave de corpo “todo verde” enquanto em Cunha (1998) lhe é atribuído o étimo tupi **ku’ruka**, “uma variedade de papagaio”. Houaiss (2001) também associa *curica* a *papagaio*, pois *curica*, segundo ele, designa “ave psitaciforme da família dos psitacídeos (*Pionopsitta caica*), da Amazônia setentrional, que atinge 23 cm de comprimento e possui plumagem verde com cabeça negra e colar amarelo de aspecto escamoso; papagainho”. Já Ferreira (2004) não faz essa associação e define *curica* como

[...] ave psitaciforme, psitacídea (*Eucinetus barrabandi*), da Amazônia, de coloração verde, cabeça, garganta e ponta da cauda pretas, faces e encontro alaranjados, peito amarelo-oliváceo, coberteiras inferiores da asa encarnadas, e rêmiges pretas marginadas de azul.

Verifica-se que a marca de regionalismo atribuída a esse item lexical, nas obras lexicográficas consultadas, foi confirmada pelos dados da pesquisa, posto que *curica* foi mencionada por um informante idoso de baixa escolaridade e uma idosa de alta escolaridade, moradores do Estado do Amapá, além do jovem de Teresina de baixa escolaridade. Ferreira (2004) e, mais especificamente Houaiss (2001), indicam a região setentrional da Amazônia como área dialetal de *curica*.

A unidade lexical *maracanã* também foi mencionada por informantes da região Norte – Macapá – como designação da “ave de penas coloridas que quando presa pode aprender a falar” (pergunta 68 do QSL/ALiB). Sampaio (1928,

p. 279) e Cunha (1998) apontam a mesma origem tupi desse item lexical, o étimo **maraká'na**, “espécie de papagaio”. Nas obras lexicográficas pesquisadas, também há consenso quanto à origem linguística do termo, uma vez que Houaiss (2001) define *maracanã* como designação de “diversas aves psitacíformes, espécie dos gêneros *Propyrrhura*, *Diopsittaca* e *Aratinga*, da família dos psitacídeos”, enquanto Ferreira (2004) acresce a essa acepção as características e as áreas de ocorrência desse gênero de aves:

[...] Bras. Zool. Designação comum às seguintes aves psitacíformes, psitacídeas: a) *Propyrrhura maracana*, distribuída por quase todo o Brasil, verde, com vértice verde-azulado, fronte escarlate, dorso inferior e meio do abdome escarlate-claros, parte basal da cauda vermelho-escura, e parte terminal e rêmiges azuis; b) *Diopsittaca nobilis*, da Amaz.; c) *Psittacara leucophthalmus*, comum em todo o Brasil, verde, com o encontro e coberteiras inferiores menores da asa encarnados.

Também *iandaia* foi mencionada como designação da ave em questão. Trata-se de uma variante de *jandaia*, segundo Cunha (1998), que registra *iandaya* como sinônimo de *jandaia*, atribuindo-lhe o étimo **ia'naia**. Já Sampaio (1928, p.245) aponta o étimo **nhand-ái** para esse termo, que nomeia “um pequeno papagaio de cabeça, peito e encontros amarelos” (SAMPAIO, 1928). Houaiss (2001), por seu turno, define a unidade lexical *jandaia* como do “tupi *ya'ndaya* ‘ave da família dos psitacídeos; espécie de papagaio”, enquanto Ferreira (2004) a define como “ave psitacíforme, psitacídea (*A. jantaya*) de coloração amarela, dorso verde, asas azuladas, cauda do verde ao azul, com ponta escura. Vive em bandos e se adapta bem ao cativeiro. Os jovens são quase totalmente verdes”. A fala do informante idoso de alta escolaridade de Cuiabá confirma o uso e, conseqüentemente, a definição fornecida por Houaiss (2001), já que ele, além de

caracterizar a *iandaia*, também afirma que *essa* ave aprende a falar⁴⁸, tipificando-a como uma espécie de papagaio.

Já a unidade lexical *mareta*, outro designativo catalogado como resposta para a pergunta 68 (QSL/ALiB), segundo Cunha (1998), deriva do étimo tupi *m̃ai'ta*, “espécie de papagaio”. Houaiss (2001), por seu turno, dicionariza *maitaca* e atribui-lhe a base tupi, *m'baita*, uma “espécie de papagaio”, definindo-a na acepção de “aves psitacíformes, da família dos psitacídeos, neotropicais, cujo corpo atarracado e cauda curta são semelhantes aos do papagaio”, tendo como sinônimos *baitaca*, *humaitá*, *maitá*, *maritaca*. A unidade lexical *maitaca* está dicionarizada na forma *maritaca* como particular da região Sudeste, na obra de Houaiss (2001), e da Amazônia (região Norte), na de Ferreira (2004). Todavia, *maritaca* foi mencionada apenas na região Centro-Oeste, pelo informante jovem do sexo masculino e de baixa escolaridade de Goiânia, capital de Goiás.

Registradas ainda na categoria das aves, vigoram *nambu* e *socó*. *nambu/nhambu* foram apontadas como respostas para a pergunta 69 do QSL (“galinha sem rabo”). Houaiss (2001), no verbete *nhambu*, remete o consulente para *inhambu*, definindo esse termo como “designação comum às aves tinamiformes, da fam. Dos tinamídeos, dos gêneros *Tinamus* e *Crypturellus*, restritas aos neotrópicos, de corpo robusto, pernas grossas e cauda rudimentar ou inexistente”. Já Cunha (1998) atribui ao verbete *inhambu* a acepção de “ave da família dos tinamídeos”, enquanto Sampaio (1928, p. 223-224) a define como “perdiz” e lhe atribui o étimo *y-nhã-bú*. Nota-se que, originalmente, o termo *nhambu* e suas variantes nomeiam uma espécie de ave com pequena ou

⁴⁸ Segue a transcrição da fala do informante, que confirma o uso e valida a resposta: “INF.- A iandaia é um papagaiozinho mais delicado, mais bonitinho, o papagaio é mais o fortão [...]. INQ. - E a iandaia aprende a falar? INF.- A iandaia é uma é uma... é uma avezinha mais assim bem trabalhada, mas também cheia, colorida tem. INQ. - E aprende a falar também? INF.- Aprende, aprende, alguma coisinha ela aprende, mas o papagaio aprende melhor”.

nenhuma cauda, e que, por extensão, passou a nomear a ave galinácea sem rabo, objeto da questão 69, ao menos para os falantes das regiões Norte e Centro-Oeste.

Também a unidade lexical *socó* foi informada como resposta para a mesma questão. Trata-se de um tupinismo originado do étimo **coó-có**, segundo Sampaio (1928, p. 315), que designa “ave que tem por hábito arrimar-se num pé só: é nome comum às pernaltas”. Já Cunha (1998) atribui à unidade lexical outro étimo (**so'ko**), acrescentando que se trata de uma “ave da família dos ardeídeos”, enquanto Ferreira (2004) define *socó* como termo da zoologia, na acepção de “designação comum a várias espécies de aves ardeídeas, especialmente as dos gêneros *Tigrisoma*, *Butorides* e *Zebrilus*. Alimentam-se de peixes e vivem ger. Isoladas ou aos pares, perto de rios, lagoas e terrenos alagadiços”.

Nota-se que nenhum dos semas atribuídos à unidade lexical *socó* aproxima-se dos traços semânticos da definição “galinha sem rabo”, ou galinha “sura”, como é designada no uso corrente, unidade léxica derivada de *suru* (variante *suro*), termo da zoologia que nomeia, segundo Ferreira (2004), o “animal sem cauda ou que só tem o coto da cauda; suri, suro, surote, bicó, cotó, nabuco, nambi, pitoco, rabi, rabicó, torado. [Fem.: *sura*.]”.

A menção dessa unidade lexical como resposta para a pergunta 69 do QSL/ALiB (“galinha sem rabo”) pode resultar do fato de o informante possuir familiaridade com o ambiente rural e, conseqüentemente, conhecer, talvez, nomes atribuídos à galinha com essa característica que possam estar caindo em desuso, uma vez que a unidade lexical foi documentada mormente na fala de idosos de alta escolaridade (um maranhense, uma amapaense e um catarinense).

Já à área semântica *mamíferos* foram vinculadas as unidades lexicais *capenga*, *caruado*, *gambá*, *jaratataca/tacaca/maria tacaca*, *mucura*, *saruê* e *timbu*. *Capenga* e *caruado* foram utilizadas pelos falantes das capitais brasileiras para nomear “o animal que tem uma perna mais curta e que puxa de uma perna” (QSL/ALiB 82). *Capenga*, que também teve documentação do gerúndio *capengando*, na base de dados do Projeto ALiB, segundo Sampaio (1928), advém do étimo tupi **cã-penga** e designa “osso torto, quebrado; o manco”. Ferreira (2004), nessa entrada, remete o consulente a coxo e Houaiss (2001) acrescenta-lhe que diz respeito a “que ou aquele que capenga, puxa da perna”, fornecendo-lhe os sinônimos: “coxo, manco, pernetá”.

Caruado, por seu turno, entendeu-se, pelo radical, ser derivado de *caruara*, aquele que sofreu dessa moléstia, unidade lexical advinda do tupi **carú-uara**, “o que come ou corróe; [...] no norte do Brasil, é uma moléstia que ataca o gado, trazendo-lhe inchação e paralyisia (sic) nas pernas e corrimentos” (SAMPAIO, 1928). Também no verbete *caruara*, Ferreira (2004) atribui-lhe a marca de uso do Nordeste e informa tratar-se do nome para a “paralisia que afeta as articulações dos bezerros e doutros animais recém-nascidos”. A marca dialetal dicionarizada em Ferreira (2004) não se confirmou nos dados analisados, haja vista que a única ocorrência de *caruado* registrou-se em Rio Branco, no Acre, aproximando o uso, porém, ao registro em Sampaio (1928), obra anterior às demais aqui utilizadas.

Dessa forma, a associação realizada pelo idoso de baixa escolaridade acreano entre o animal manco e o nome *caruado* pode residir no fato de ele raciocinar que esse animal sofreu de caruara e, por isso, ficou coxo, o que valida a resposta como nome para o animal manco.

E ainda designando mamíferos, foram registradas as unidades léxicas de origem tupi *gambá*, *jaratataca/tacaca/maria tacaca*, *mucura*, *saruê* e *timbu*, todas fornecidas como resposta para a questão 71 do Questionário Semântico-Lexical do Projeto ALiB (“animal que solta um cheiro ruim quando se sente ameaçado”).

A unidade lexical *gambá*, inclusive, parece constituir-se em norma lexical nacional para designar o animal em questão, uma vez que foi mencionada pelos habitantes das 25 capitais pesquisadas entrevistados pela equipe de pesquisa do Projeto ALiB, inclusive na região Norte, demonstrando que essa unidade lexical representa uma forma inovadora no Norte do Brasil, provavelmente em decorrência dos processos migratórios.

Sampaio (1928, p. 201) define *gambá* como “o ventre aberto, a barriga ôca (sic)”, uma clara referência ao marsúpio, a bolsa abdominal das fêmeas desse animal onde os recém-nascidos se alojam e onde se encontram também as tetas. Trata-se de uma unidade léxica originada do tupi **guá’mbá** que, segundo Cunha (1998), é o “nome comum a vários mamíferos marsupiais do gênero *Didelphis*”. Houaiss (2001) acrescenta a essa definição que os marsupiais são “encontrados do Sul do Canadá à Argentina, com até 50 cm de comprimento, cauda preênsil, longa e quase inteiramente nua, com a parte distal branca, pelagem cinza, preta ou avermelhada e fêmeas com marsúpio bem desenvolvido”.

Comparando-se as definições, percebe-se uma ampliação de sentido desse item lexical, pois, enquanto em Sampaio (1928, p. 201) é definido como o ventre aberto característico de uma espécie animal, em Houaiss (2001) já

aparece como designativa do próprio animal, a mesma acepção com que foi documentada pelos pesquisadores do Projeto ALiB.

A pesquisa etimológica sobre a unidade lexical *mucura*, por seu turno, evidenciou posições controversas quanto à base linguística que lhe deu origem. Todavia, neste trabalho, foi aceita a tese da sua origem tupi, pois as duas obras que tratam de indigenismos e que embasaram este estudo atribuem-lhe tal origem. Sampaio (1928, p. 287), por exemplo, registra *mucura* como do étimo tupi **mô'cura**, que indica aquele que “faz que absorva ou que se oculte dentro de si mesmo; alusão ao fato de esconder o animal deste nome os próprios filhos no saco ou bolsa que tem no ventre. É o marsúpio, também conhecido como gambá ou sariguê”. Cunha (1998), por sua vez, atribui a esse item lexical o étimo **mu'kura**, provavelmente derivado de *mô'cura*, uma vez que a acepção é essencialmente a mesma, “mamífero marsupial da família dos didelfídeos; gambá, sariguê”. Houaiss (2001) e Ferreira (2004) também associam a unidade lexical *mucura* a *gambá*, posto que, em ambas as obras, no verbete *mucura*, os lexicógrafos ratificam a sua origem tupi, fazendo remissão ao verbete *gambá*.

Todavia, há alguns estudiosos que contestam a origem tupi do vocábulo *mucura*. Cardoso (1961, p. 203), por exemplo, aponta-lhe o étimo macuxí **mocorê**, para nomear os “didelfos”. Apesar de outros renomados estudiosos do léxico imputarem à forma a origem tupi, como Nascentes (1966) e Tibiriçá (1997), a base macuxí, advinda do dialeto da família Karib, atribuída por Cardoso, deve ser fonte para estudos e reflexões, por ser o autor reconhecido como um importante amerigenista, investigador da influência indígena no português do Brasil, especialmente da tupi-guarani.

Além disso, Machado (1987) também hesita ao registrar a origem desse termo, à medida que considera *mucura* como advinda “talvez do tupi mo (faz) + cur (ser tragado)”. Há ainda uma terceira hipótese aventada para a etimologia dessa unidade lexical, apresentada por José Gregório (1980, p. 955), estudioso das línguas indígenas brasileiras que aponta a língua guarani, do tronco tupi, como a origem desse item lexical. Pelo exposto, nota-se que a questão da etimologia da unidade lexical *mucura* carece de estudos mais aprofundados, tendo neste trabalho sido considerado, por ora, um tupinismo, com base nas ponderações já expostas.

Ainda no rol de unidades lexicais de origem tupi documentadas como respostas para a pergunta 71 do QSL/ALiB, temos o item léxico *saruê*, que designa o “bicho que solta cheiro ruim quando se sente ameaçado”. Nesse caso, todos os estudos apontam a língua tupi como fonte originária, como em Sampaio (1928, p. 313), onde se registra o étimo **çooó-r-ighê**, que significa “o animal de saco”, ou em Cunha (1998), que lhe atribui o étimo **sari'üe** na acepção de “mamífero marsupial da família dos didelfídeos; gambá (sarôhê)”. Também Houaiss (2001) e Ferreira (2004) apontam para *saruê* a acepção de *gambá*. Ferreira (2004) classifica *saruê* nessa acepção como um brasileirismo do Nordeste. Todavia, nos dados levantados pelo Projeto ALiB, houve seu registro apenas na região Sudeste (Vitória-ES), o que denota um indício de processo de mudança quanto ao uso do termo e ratifica a importância das pesquisas geolinguísticas para a documentação de mudanças em curso na língua.

Também foram documentados *jaratataca/tacaca* e *maria tacaca* como designações para o mesmo mamífero, aqui consideradas como variantes.

Segundo Houaiss (2001), a *jaritataca* (para onde a consulta ao verbete *jaratataca* remete) é um

[...] mamífero carnívoro, mustelídeo (*Conepatus chilensis amazonicus*), amplamente distribuído no Norte do Brasil, de coloração preta, com uma faixa branca dorsal dividida longitudinalmente ao meio por uma linha preta. Mede 50cm além da cauda, que só tem 15cm e é muito pilosa, branca e empenachada; é provido de uma glândula anal que segrega e faz projetar, como defesa, um líquido fétido, irritante e nauseante.

Já de acordo com Ferreira (2004), *tacaca* deriva de *maritacaca*, por aférese, processo de supressão de um fonema ou grupo de fonemas no início de uma palavra. A análise semântico-lexical dessa unidade lexical também apontou incerteza, não relacionada à origem, como no caso de *mucura*, mas em relação ao conceito que designa, posto que, em Cunha (1998), o vocábulo de étimo tupi **mjiara'taka** é definido como o “mamífero carnívoro da família dos mustelídeos”. Sampaio (1928), por seu turno, não faz referência a esses termos, enquanto Ferreira (2004), no verbete *tacaca*, registra a remissão a *gambá*, como uma das acepções do termo, atribuindo-lhe a marca dialetal de brasileirismo do Nordeste, marca confirmada pelos dados do Projeto ALiB, uma vez que essa unidade léxica só foi documentada em duas capitais da região Nordeste – João Pessoa e Natal.

Também o item lexical *timbu* foi informado pelos falantes das regiões Norte e Nordeste como designação do “bicho que solta cheiro ruim quando se sente ameaçado”. O étimo que o originou, apesar de haver consenso de que seja tupi, é incerto: na obra de Sampaio (1928) não há registro. Cunha (1998), apesar de relacioná-lo como tupinismo na acepção de “marsupial da família dos didelfídeos; espécie de gambá”, não faz referência à sua etimologia, documentada apenas em Houaiss (2001) – *taibu* – que designa o “gambá”. Ferreira (2004) o registra como de “possível” origem tupi, classificando-o como

brasileirismo do Nordeste e remetendo o consulente para o verbete *gambá*. Essa marca de uso foi parcialmente confirmada pelos dados do ALiB, haja vista que o termo *timbu* foi citado por dois natalenses e por uma recifense; porém, também um portovelhense o mencionou como resposta para a pergunta 71 (QSL/ALiB).

Vê-se assim que, das cinco unidades lexicais registradas como designativos para o “animal que solta um cheiro ruim quando se sente ameaçado”, quatro delas são dicionarizadas como sinônimas: *gambá*, *mucura*, *saruê* e *timbu*; apenas *jaratataca/variantes* têm definição diferente, tratando-se de outro referente. Todavia, se atentarmos para a dicionarização de *gambá* e de *jaratataca*, notamos que o animal que, de fato, corresponde ao conceito expresso na pergunta 71, (QSL/ALiB) é o mustelídeo e não o marsupial. Para uma melhor visualização dessa realidade, elaborou-se o Quadro XI, a seguir, que visualiza as informações relativas à dicionarização dos itens lexicais em causa, comparando-as:

| Unidade Lexical | Família (Zool.) | Tamanho | Características físicas | Pelagem | Alimentação | Odor fétido |
|-------------------|-----------------|----------------|-------------------------|--|--|-------------------------------|
| Gambá | Didelfídeo | Cerca de 50 cm | Cauda longa e quase nua | Negra, cinza ou avermelhada, com a parte distal branca | ovos, frutos, vermes, insetos, lagartos, anfíbios e até mesmo filhotes de pássaros | Não relacionado |
| Jaratataca | Mustelídeo | Cerca de 40 cm | Cauda branca | Negra ou marrom escura, com duas faixas brancas | insetos, pequenos vertebrados e frutos | Exalado pelas glândulas anais |

Quadro XI: Traços semânticos das unidades lexicais *gambá* e *jaratataca*

Os traços semânticos contidos no quadro confirmam que as unidades lexicais *gambá* e *jaratataca* nomeiam referentes distintos, apesar de as características de ambos serem bastante similares. Nesse contexto, uma primeira avaliação descartaria a unidade léxica *jaratataca/arataca* como nominativo do

animal normalmente designado como *gambá*. Porém, se considerados os semas requeridos pela pergunta, nota-se que a resposta é condizente com a questão 71 do QSL/ALiB e, conseqüentemente, justifica a inclusão dessa unidade lexical nos dados analisados neste trabalho, que examina os nomes para o “bicho que solta cheiro ruim quando se sente ameaçado”. Além disso, vê-se que a resposta inicialmente esperada para a pergunta 71, *gambá*, não exala odor fétido quando ameaçado, apenas a jaratataca, podendo esse traço semântico lhe ter sido associado apenas culturalmente, não se confirmando, porém.

As unidades lexicais registradas como resposta para a pergunta em questão (QSL/ALiB), inclusive, foram tema também de outro estudo geolinguístico realizado por Silva e Aguilera (2010), que analisou todas as unidades lexicais registradas como respostas para a questão 71 do Questionário Semântico-Lexical do Projeto ALiB, não apenas as de base indígena. Esse trabalho, dentre outras, teve como algumas de suas conclusões, que “(i) a maior parte das variantes registradas pelos falantes são provenientes do tupi, fato este que demonstra a influência indígena no campo lexical da fauna brasileira; (ii) a variante *gambá* é predominante em praticamente todo o território analisado” (SILVA; AGUILERA, 2010, conclusões essas que condizem também com os resultados deste estudo⁴⁹).

Já o grupo relacionado aos *insetos* abrigou, neste estudo, sete itens lexicais de base tupi: *carapanã*, *cupim*, *maruim/muruim*, *muriçoca*, *mutuca/motuca*, *tapuru/taparutapuruzim* e *туру/toru*. O primeiro foi documentado como resposta para a pergunta 88 do Questionário Semântico-Lexical (ALiB) que busca designações para o “inseto pequeno de pernas cumpridas, que canta no ouvido das pessoas à noite”. Existe uma discussão teórica acerca da origem

⁴⁹ Disponível em: <http://www.cielli.com.br/downloads/529.pdf>),

dessa unidade lexical, porém, neste trabalho, foi adotada a mesma postura tomada no caso de *mucura* (QSL/ALiB 071), considerando-se que, no universo de obras pesquisadas, a maioria indicou a etimologia tupi para *carapanã*, por isso essa unidade lexical foi aqui classificada como tupinismo, não se descartando, obviamente, a necessidade de estudos mais aprofundados acerca de base linguística dessa palavra.

Carapanã, segundo Cunha (1998), deriva do étimo **karapa'na** e nomeia uma “espécie de mosquitos da família dos culicídeos; na Amazônia, designação genérica de mosquito”. Já em Sampaio (1928) não se encontra essa unidade lexical. Ferreira (2004), por sua vez, classifica-a como brasileirismo amazônico na acepção de mosquito:

[...] designação comum a todos os insetos dípteros, culicídeos, de porte pequeno, pernas muito longas, corpo e asas revestidos de escamas, antenas longas e finas, com 16 artículos. O ciclo evolutivo efetua-se em duas fases distintas: a primeira, na água, onde são depositados os ovos e se desenvolvem as larvas e pupas; a segunda é alada e terrestre. [Sin.: *pernilongo*, *mosquito-prego*, *muriçoca*, *carapanã*, *carapanã-pinima*, *fincão*, *fincudo*, *sovela*, *perereca*, *bicuda*].

Machado (1987), por seu turno, aponta que o termo origina-se do “tupi **cará** (ave) + **panã** (que corta, morde, pica)” e Nascentes (1955) concorda com o étimo atribuído por Cunha (1998) a *carapanã*. Já Cardoso (1961, p. 114), contrariando a posição desses dicionaristas, confere a essa unidade léxica a etimologia **galibí**, do tronco caribe, *karapanã* ou *carapanã*. Tibiriçá (1997), em seu *Dicionário de topônimos brasileiros de origem tupi. Significação de nomes geográficos de origem tupi*, também atribui a origem **galibí** ao termo, esclarecendo, porém, que esse étimo entrou em contato com o português “através do tupi”, tratando-se do “nome comum do mosquito, no norte do Brasil”. Neste estudo, a unidade lexical *carapanã* também foi considerada um tupinismo,

com base na obra de Cunha (1998), referencial para esta pesquisa. Os dados do Projeto ALiB confirmam a marca de regionalismo atribuída ao termo por Ferreira (2004), uma vez a documentação de *carapanã* restringiu-se às capitais da região Norte.

Ainda foram fornecidas como respostas para a questão 88 (“inseto pequeno de pernas cumpridas, que canta no ouvido das pessoas à noite”) a unidade lexical *maruim* (e sua variante *murui*) e *muriçoca*. No caso da primeira, foram aqui tomadas como variantes, haja vista que *maruim* e *murui* são apontadas como sinônimos nas obras de Houaiss (2001) e de Ferreira (2004). *Maroim*, segundo Sampaio (1928, p. 281), deriva de **mberuí** e designa “os mosquitos”. Já para Cunha (1998), deriva do étimo **mari’üi**, também tupi, o “nome comum a vários mosquitos hematófagos”. Houaiss (2001), por sua vez, define *maruim* como

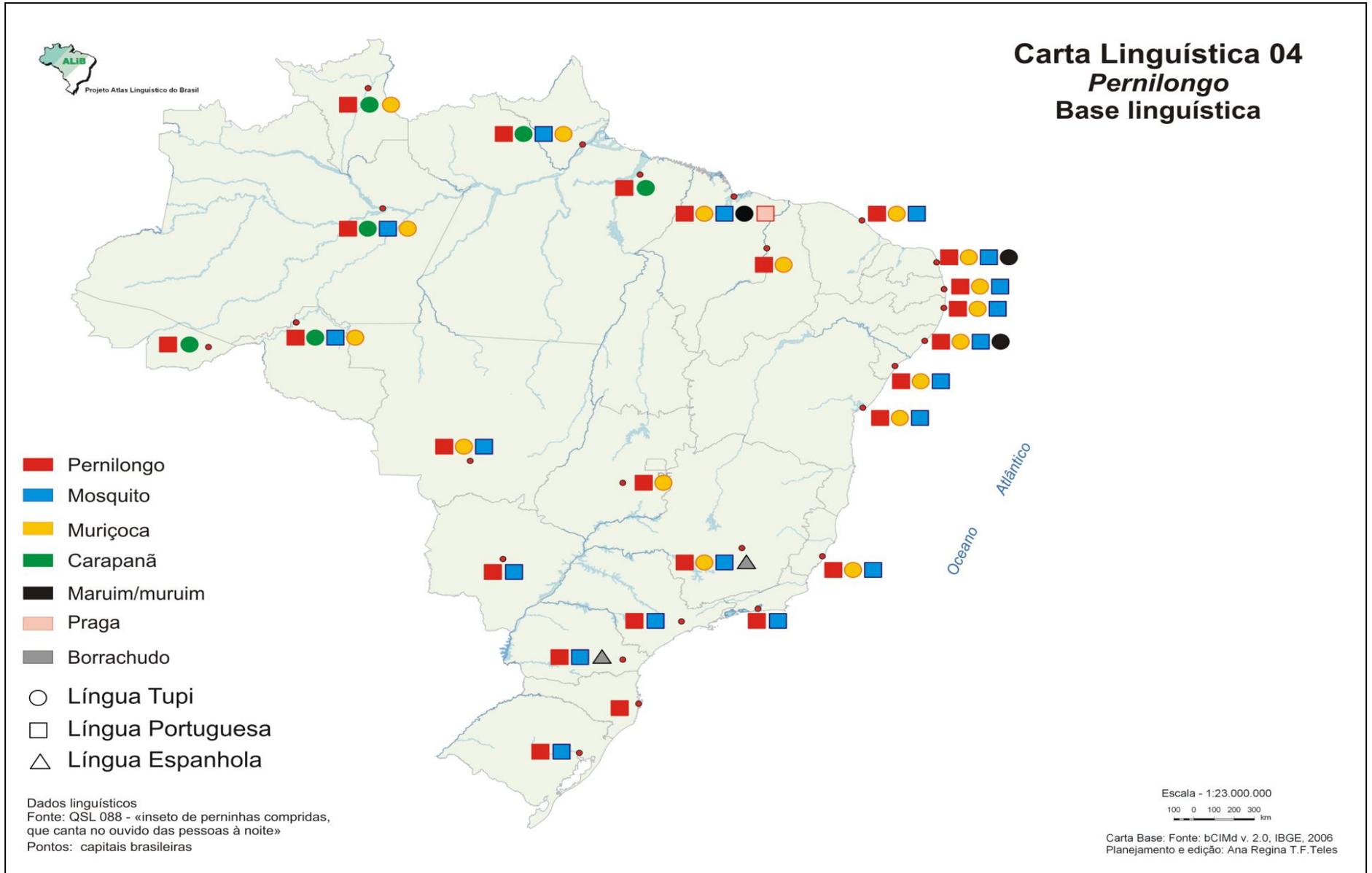
[...] designação comum a diversos mosquitos da família dos ceratopogonídeos, de até 2 mm de comprimento, cuja ocorrência no Brasil está associada aos manguezais; as fêmeas são hematófagas e transmissoras da filariose ao homem e aos animais domésticos por meio de picadas dolorosas.

Os dados coletados pelo Projeto ALiB e analisados no âmbito deste estudo confirmam a informação fornecida por Houaiss (2001), que documenta a ocorrência do inseto *murui* em manguezais, ecossistema encontrado no Nordeste brasileiro, região onde se documentou essa designação: Maceió (01 ocorrência) e São Luiz (02 registros) e Natal (01 ocorrência). A informante idosa com Ensino Fundamental de Maceió, inclusive, associa a presença desse inseto às áreas alagadas, pois informa que o *maruim* é o inseto “pequeninho que dá na beira da lagoa”.

Já *muriçoca*, outro tupinismo que nomeia o “inseto pequeno de pernas cumpridas, que canta no ouvido das pessoas à noite” (QSL/ALiB 88), provém do étimo tupi **muri’soka**, sendo definida por Cunha (1998) como “variedade de mosquito”. Houaiss (2001), por sua vez, no verbete *muriçoca*, remete o consulente para mosquito, definindo esse termo como “designação comum aos insetos dípteros, de pequeno tamanho, especialmente os hematófagos da família dos culicídeos, gerando vetores de conhecidas doenças do homem”, unidade léxica sinônima de *bicuda*, *carapanã* e *pernilongo*, dentre outras. Essa unidade lexical ocorreu nas respostas de falantes de todas as capitais da região Nordeste, tendo também se registrado na região Norte, onde *carapanã*, contudo, reina em absoluto em todas as capitais, tendo sido citada por 100% dos informantes, além de ter documentação também na região Centro-Oeste e Sudeste, ausentando-se apenas no Sul, como a continuidade do trabalho demonstrará.

Nas obras lexicográficas consultadas, *muriçoca* aparece classificada com marca regional do Nordeste e do Estado de Minas Gerais. Todavia, os dados do ALiB demonstram que a disseminação desse item lexical está para além do dicionarizado, o que pode ser atribuído às migrações de nordestinos para todas as regiões do Brasil.

Assim como ocorreu com os designativos para *gambá*, também os nomes para o “inseto pequeno de pernas cumpridas, que canta no ouvido das pessoas à noite” (QSL/ALiB 88) são em sua grande maioria de base indígena, neste caso também a tupi, uma vez que 03 das 07 unidades lexicais registradas são dessa origem. A Carta Linguística a seguir demonstra o exposto, além de visualizar a distribuição diatópica das unidades léxicas elencadas como designação desse conceito:



Vê-se assim que a língua tupi divide com a portuguesa a liderança nos registros de base linguística dos designativos para pernilongo, confirmando que a herança indígena se faz presente nos nomes para referentes ligados à fauna brasileira, como o decorrer desta pesquisa tem demonstrado.

Ainda na área semântica *insetos*, a questão 83 do questionário semântico lexical do Projeto ALiB busca nomes para “um tipo de mosca grande, esverdeada, que faz barulho quando voa”. Das respostas fornecidas pelos falantes, foram documentados dois tupinismos, *mutuca* e *motuca*, que foram neste estudo consideradas variantes, embasando-se na sinonímia aplicada a esses vocábulos em Houaiss (2001) e em Ferreira (2004). Sampaio (1928, p. 287) atribui-lhe o étimo tupi **mô-tuca** (“faz que perfure”). Já na obra de Cunha (1998), aponta-se o étimo **mu'tuka**, “nome comum às moscas da família dos tabanídeos”. Ferreira (2004), por sua vez, define *mutuca* como “designação comum aos dípteros da família dos tabanídeos” e, a essa acepção, Houaiss (2001) acrescenta que é “de corpo robusto e de tamanho médio a grande, sendo apenas as fêmeas hematófagas; butuca, moscardo, motuca, tavão [São incômodas ao gado e ao homem, devido às suas picadas dolorosas]”.

Mutuca foi considerada válida como resposta para o conceito em pauta haja vista, em primeiro lugar, sua alta produtividade (14 ocorrências), além da argumentação exposta pelos informantes entrevistados. O jovem manauara de alta escolaridade informou que *mutuca* é o nome comum na região para esse tipo de mosca, o que se confirma por esse item lexical haver sido mencionado por dois manauaras e dois amapaenses, além de dois paraenses, um piauiense, dois paraibanos e um maranhense, dentre os demais registros⁵⁰.

⁵⁰ Também mencionaram *mutuca* como resposta um cuiabano, um natalense, um aracajuense e um morador de Florianópolis.

Como resposta para a questão 87 do QSL/ALiB, que busca nomes para “bicho que dá em esterco, pau podre”, vigoram três tupinismos considerados como respostas válidas – *cupim*, *tapuru/taparu/tapuruzim* e *туру/tору*. Nas obras pesquisadas, não está apontada essa acepção para a unidade léxica *cupim*, que é oriunda do tupi **kupi'i** e designa os “insetos da ordem dos isópteros” (CUNHA, 1998) ou *copim*, “formiga branca” (SAMPAIO, 1924, p.191). A obra de Ferreira (2004) também define *cupim* como “designação comum aos insetos isópteros. [...] [que] constroem cupinzeiros na madeira ou no solo. [...] mas podem alimentar-se, também, de objetos de madeira ou compensado, de papel, etc”, enquanto Houaiss (2001) aponta *tapuru* como sinônimo de *cupim*.

Os dados do ALiB confirmam o uso dessa variante lexical como o dicionarizado em Houaiss (2001), dado que os falantes de todas as regiões político-geográficas brasileiras utilizam *cupim* para nomear “bicho que dá em esterco, pau podre”, cujas respostas poderiam ser também *coró*, *bicho*, etc, contabilizando 29 ocorrências na base de dados do ALiB. No verbete *tapuru*, inclusive, o lexicógrafo registra a seguinte informação: “regionalismo de Mato Grosso, mesmo que *cupim* (‘designação comum’)” (HOUAISS, 2001).

Assim, dada a alta produtividade do designativo *cupim* para o referente em questão, considerou-se essa uma resposta válida. Além disso, também o *habitat* das duas espécies se aproxima, no caso a madeira, o que relaciona *cupim* aos semas requeridos pela pergunta 87 do Questionário Semântico-Lexical do Projeto Atlas Linguístico do Brasil.

Também vinculada ao grupo dos *insetos*, houve o tupinismo *tapuru* e de suas variantes *taparu/tapuruzim*, como designação para o conceito a que se refere à questão 87 (QSL/ALiB) (bicho de esterco), assim como na questão 86,

que busca nomes para “aquele bichinho branco, enrugadinho, que dá em goiaba, em coco”. Cunha (1998) define *tapuru* como “espécie de larva vermiforme de insetos dípteros, que depositam os ovos nas carnes em putrefação; bicheira”, originado do étimo tupi **tapu’ru**. Já Houaiss (2001) atribui, para esse item, as acepções “bicheira” e “cupim”, enquanto Ferreira (2004), além de defini-lo como “bicheira”, classifica-o como sinônimo de “bicho da fruta”, o que pode justificar o uso de *tapuru* por habitantes das regiões Norte e Nordeste para designar o “bicho da goiaba”.

Para Houaiss (2001), *tapuru* é um regionalismo de Mato Grosso, enquanto Ferreira (2004) classifica-o como brasileirismo do Nordeste. Essa unidade lexical, no falar do homem citadino apenas das regiões Norte e Nordeste do Brasil, confirma em parte a marca dialetal a ela atribuída por Ferreira (2004), pois foi documentada em seis capitais nordestinas – Maceió, João Pessoa, Recife, Natal, Fortaleza e São Luiz. Além disso, os dados do Projeto ALiB demonstram a produtividade da unidade léxica *tapuru* em todas as capitais da região Norte.

Finalizando, houve também o registro, como resposta para a questão 87 (QSL/ALiB) (bicho de esterco, pau podre), das unidades lexicais de origem tupi *туру* e *тору*. Sampaio (1928, p. 335) define *туру* como “nome dado geralmente a animais aquáticos e vermes”, enquanto Cunha (1998) a define como “molusco bivalve da família dos teredinídeos, de aspecto vermiforme” sem, contudo, atribuir-lhe um étimo. Houaiss (2001), no verbete *туру*, remete o consulente a *gusano*, “designação comum aos moluscos bivalves vermiformes, [que] [cavam galerias em madeira submersa, o que constitui um sério problema para os cais e embarcações construídos com madeira]”. Podem-se associar, destarte, a partir do

sema “cavar galerias em madeiras submersas”, a pergunta do QSL/ALiB e a documentação na fala dos falantes e o dicionarizado na obra de Houaiss (2001). Além disso, percebe-se uma marca dialetal no uso de *turu* para designar o bicho de esterco, haja vista que 04 moradores do Amapá e mais 04 de Belém mencionaram esse item lexical para designar o conceito em questão.

Concluindo-se assim a análise léxico-semântica das unidades lexicais auferidas como respostas para as perguntas pertinentes à *fauna* (QSL/ALiB), segue-se agora a análise no tocante a sua distribuição diatópica. Para tanto, é necessário situar inicialmente os itens lexicais no espaço em que foram documentados, o que é visualizado pelo Quadro XI a seguir:

| QSL/pergunta | Tupinismo | N | NE | CO | SE | S |
|---|----------------------------------|---|----|----|----|---|
| 64. “ave preta que come animal morto, podre” | urubu | X | X | X | X | X |
| 68. “ave de penas coloridas que quando presa pode aprender a falar” | curica | X | X | | | |
| | maracanã | X | | | | |
| | arara | X | X | X | X | X |
| | maretaça | | | X | | |
| | iandaia | | | X | | |
| | ajuru | X | | | | |
| 71. “bicho que solta cheiro ruim quando se sente ameaçado” | gambá | X | X | X | X | X |
| | mucura | X | X | | | |
| | timbu | X | X | | | |
| | saruê | | | | X | |
| | jaratataca/tacaca / maria tacaca | | X | X | | |
| 82. “o animal que tem uma perna mais curta e que puxa de uma perna” | capenga | X | X | | X | X |
| | caruado | X | | | | |

| | | | | | | |
|---|-----------------------------|---|---|---|---|---|
| 83. “um tipo de mosca grande, esverdeada, que faz barulho quando voa” | mutuca / motuca | X | X | X | | X |
| 86. “aquele bichinho branco, enrugadinho, que dá em goiaba, em coco” | tapuru / taparu / tapuruzim | X | X | | | |
| 87. “bicho que dá em esterco ou pau podre” | tapuru / taparu / tapuruzim | X | X | X | | |
| | turu / toru | X | | | | |
| | cupim | X | X | X | X | X |
| 88. “inseto pequeno de pernas cumpridas, que canta no ouvido das pessoas à noite” | carapanã | X | | | | |
| | muriçoca | X | X | X | X | |
| | maruim / muruim | | X | | | |

Quadro XII – Distribuição dos tupinismos na área semântica da *fauna* documentados nas capitais brasileiras, de acordo com a região do Brasil.

Ao analisarmos os dados documentados pelo Projeto ALiB nas capitais brasileiras selecionados para este estudo, sob o ponto de vista diatópico, notamos que os falantes das regiões político-geográficas do Brasil portam-se de formas distintas quanto ao uso de itens lexicais de base indígena. Na região Norte, por exemplo, registrou-se o maior número de tupinismos (16): *urubu*, *curica*, *maracanã*, *arara*, *ajuru*, *gambá*, *mucura*, *timbu*, *capenga*, *caruado*, *mutuca/motuca*, *tapuru/taparu/tapuruzim*, *turu/toru*, *cupim*, *carapanã* e *muriçoca*. Algumas motivações histórico-geográfico-econômicas podem justificar esse comportamento linguístico dos habitantes da região amazônica.

A região Norte ainda concentra a maior população indígena do País – aproximadamente 213 mil indivíduos (IBGE, 2010) –, fato já investigado por estudiosos brasileiros, como Prado Jr. que, em seu livro *Formação do Brasil contemporâneo*, de 1942 (1ª edição), apontou que há algumas “regiões [onde] o sangue indígena prepondera. São em geral as de vida econômica pouco ativa, ou

onde a colonização não tomou pé muito acentuado” (PRADO JR, 1986, p. 113). Isso poderia justificar a maior produtividade de tupinismos no vocabulário dos habitantes das capitais da região Norte. Também deve ser considerado nesse contexto o *nheengatu*⁵¹, língua que possui semelhanças com o tupi e ainda em uso por caboclos ribeirinhos em diversas comunidades, dentre outras, São Gabriel da Cachoeira, no Amazonas, “onde em plena cidade, em quase todas as residências de gente civilizada, se fala corretamente a língua *nheengatu*” (IBGE, 2009). Além disso, sempre houve informações de uma maior concentração dos índios tupi em áreas de grandes rios e florestas, pois se destacavam como grandes remadores e talentosos na lida extrativista, tanto no início da colonização quanto mesmo nos dias atuais.

Já a região Nordeste ocupou o segundo lugar em termos de documentação de tupinismos no âmbito do universo aqui pesquisado (13): *urubu, curica, arara, gambá, mucura, timbu, jaratataca/tacaca/Maria tacaca, capenga, mutuca/motuca, tapuru/taparu/tapuruzim, cupim, muriçoca e maruim/muruim*. Podem justificar esse comportamento linguístico tanto a proximidade com a região Norte quanto, especialmente, o caráter conservador documentado na fala dos nordestinos, pois “a área nordestina é, talvez, juntamente com a da Amazônia, em que mais vulgarizados estão, na língua do povo, os vocábulos arcaicos” (CARDOSO, 1961, p. 88). Além disso, há que se considerar a presença de indígenas no Nordeste. Atualmente, existem ainda os Potiguara, na Paraíba, os Pankaru, no Pernambuco e os Pataxó, na Bahia, (MEADER. Disponível em: <http://www.sil.org/americas/brasil/publcons/ling/indnord.pdf>). As duas primeiras

⁵¹ À língua geral amazônica, produto de interação entre os portugueses e tupinambás no século XVII, associou-se o nome *nheengatu* a partir do final do século XIX, e “seu uso se estendeu pelo Solimões até a Amazônia peruana e pelo Rio Negro até a Venezuela e a Colômbia” (RODRIGUES, 2002, p. 45).

etnias, inclusive, são falantes de línguas do tronco Tupi (a língua pataxó é pertencente ao tronco linguístico Macro-Jê).

Na sequência, em relação ao número de variantes lexicais de base tupi selecionadas para este estudo, aparece a região Centro-Oeste, com 10 ocorrências: *urubu*, *arara*, *maretaca*, *iandaia*, *gambá*, *jaratataca/tacaca/maria tacaca*, *mutuca*, *tapuru/taparu/tapuruzim*, *cupim* e *muriçoca*. Esses dados podem se justificar dadas duas características principais dessa região: se, por um lado, ela agrega a segunda maior comunidade indígena no Brasil (IBGE, 2010), por outro, é marcada também pelo alto processo migratório de diversas outras etnias e povos, como paraguaios, libaneses, japoneses, além de mineiros, paulistas e sulistas.

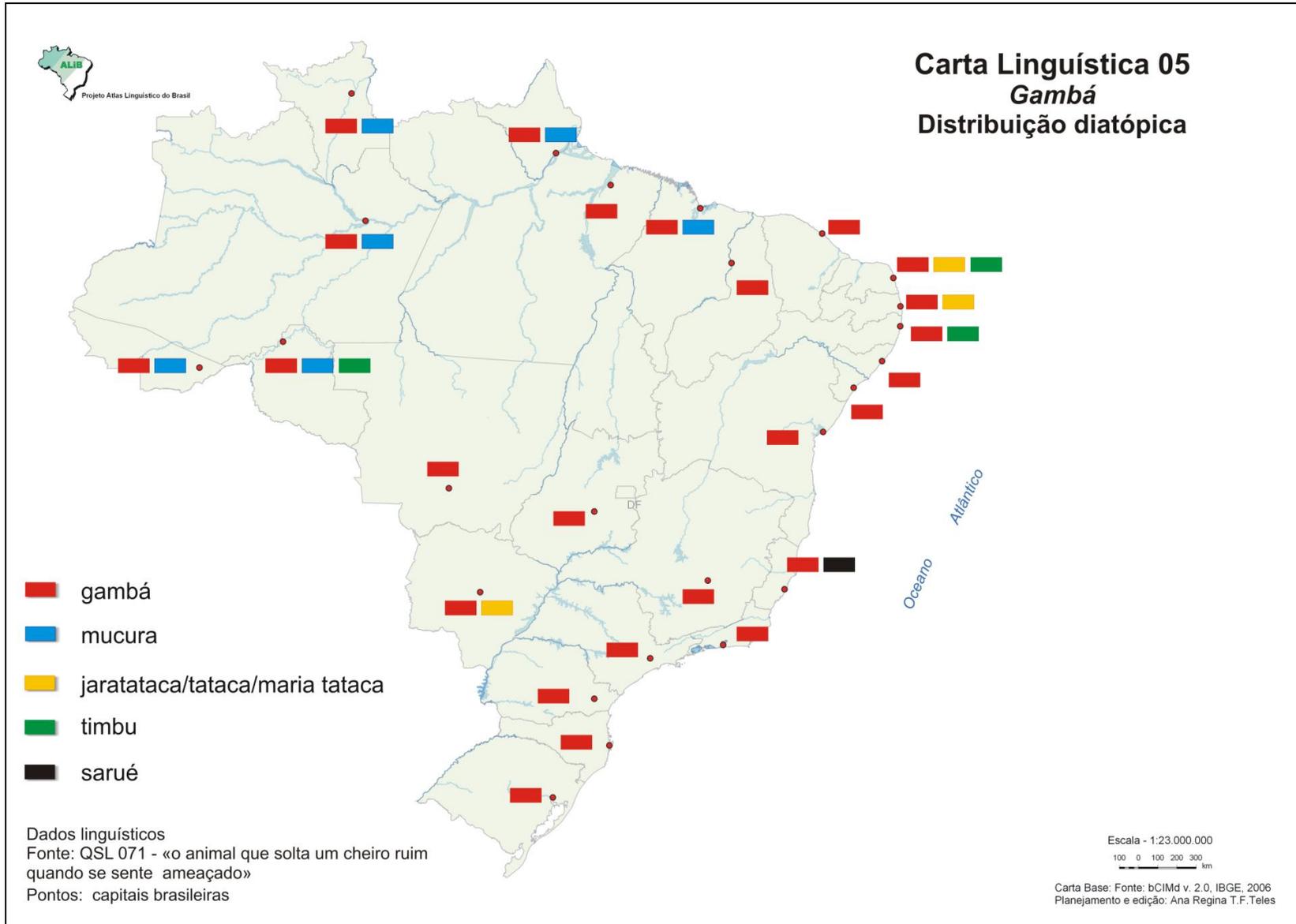
Depois, com o menor registro de indigenismos documentados, encontram-se as regiões Sudeste, com 07 registros: *urubu*, *arara*, *gambá*, *saruê*, *capenga*, *cupim* e *muriçoca* e Sul, com 06 indigenismos: *urubu*, *arara*, *gambá*, *capenga*, *mutuca/motuca* e *cupim*. A baixa presença no sul do Brasil pode dever-se ao fato de essa região caracterizar-se pela presença marcante de imigrantes não lusos em sua base cultural e populacional, como a alemã, italiana, polonesa, pomerana no Sul, assim como poucas são as comunidades indígenas no local.

Já o Sudeste pode também ter sua baixa produtividade de indigenismos associada a questões extralinguísticas, uma vez que, via lavouras de café, concentrou renda capaz de arcar com a aquisição de escravos africanos no século XVIII e início do século XIX, como o assinalado por Prado Jr (1986, p.112): “o preto e seus derivados dominam nas regiões de grande atividade econômica, presente ou passada (escravo e trabalhador)”. E essa mão de obra foi, depois da Abolição da Escravatura (1888), substituída por imigrantes de

outros continentes, especialmente pelos italianos, que se firmaram na região e também deixaram suas marcas na cultura e no léxico dos habitantes dessa região. Além disso, o Sudeste concentra a maior densidade demográfica do Brasil, pois, desde o século XIX, apresenta um intenso processo de urbanização e industrialização, atraindo imigrantes europeus e asiáticos, além de ser polo turístico internacional e detentor de megalópoles de importância mundial, como São Paulo, capital brasileira em que se documentou o menor número de indigenismos (03: *urubu, gambá e cupim*).

Porém, o Sudeste também pode manter essas unidades léxicas em uso em decorrência da própria história social da região, marcada também pelas bandeiras e cuja influência linguística indígena permanece ainda arraigada naquele território, como a grande produtividade de topônimos de origem indígena em São Paulo e Rio de Janeiro, por exemplo. Conforme Dick (1994, p. 435), os topônimos indígenas, especialmente os tupis, têm sua frequência registrada em índices expressivos na norma lexical brasileira “nos mais variados itens semânticos (a exemplo de zoonímia, fitonímia, hidronímia, geomorfonímia, ergononímia)”.

Dentre os indigenismos documentados pela equipe de pesquisadores do Projeto ALiB na área semântica da *fauna*, percebe-se que alguns deles podem demarcar áreas dialetais. É o caso, por exemplo, da unidade léxica *mucura*, para designar o animal também conhecido como *gambá*. Para visualizar essa realidade, elaborou-se a Carta Linguística V, a seguir:



Nota-se, pela análise da Carta Linguística V, que *gambá* está disseminado por todas as capitais brasileiras, e as demais designações documentaram-se de maneira irregular em território nacional, à exceção de *mucura*, que desenha um *continuum* geográfico de ocorrência dominando o Norte do país e se estendendo ao Estado nordestino vizinho, o Maranhão.

E esse *continuum* geográfico fica ainda mais evidente se forem considerados os percentuais de ocorrência dessa unidade lexical nas capitais do Norte e Nordeste do país. Relembrando a dicionarização da marca de uso do termo, Houaiss (2001) atribui-lhe a marca de regionalismo da Amazônia, enquanto Ferreira (2004) não lhe imputa nenhuma referência dialetal. De fato, os registros de *mucura* aproximam-se da dicionarização em Houaiss (2001), tendo em vista que a maior ocorrência do designativo em pauta foi, de fato, na região amazônica e, afastando-se dessa área, sua frequência vai diminuindo, como se pode ver a partir do comparativo entre o mapa da Carta Linguística V e o Gráfico VI a seguir:

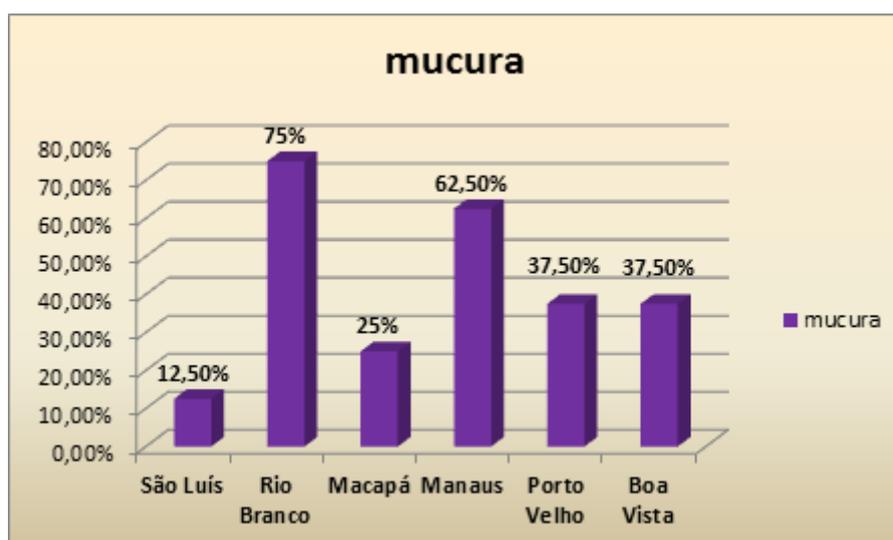


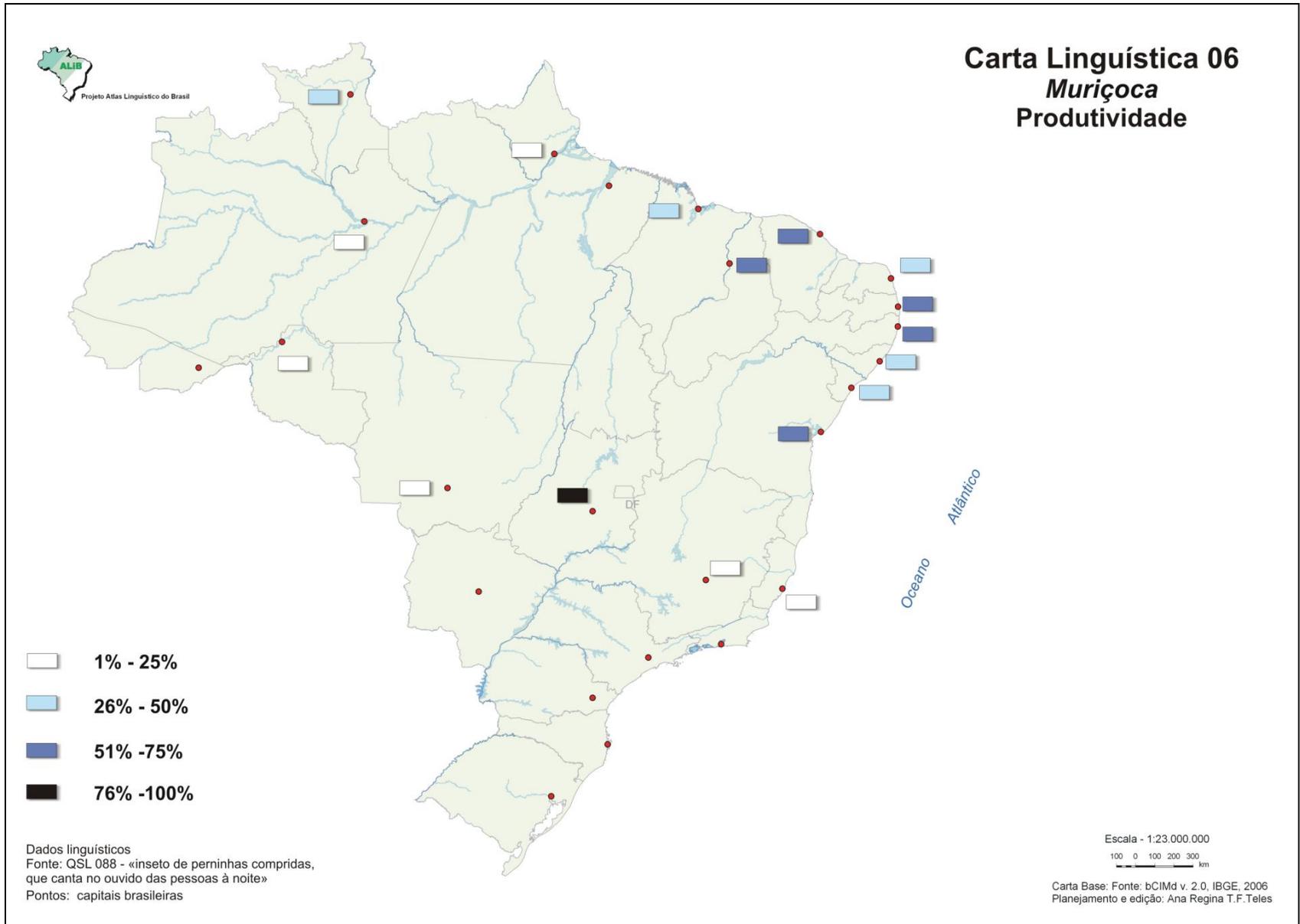
Gráfico VI – Produtividade de *mucura* nas capitais brasileiras

Considerando o universo de oito informantes por capital pesquisada, vê-se que o menor registro foi na capital nordestina de São Luís (12,5%).

Seguindo pela região amazônica, o maior percentual de ocorrência deu-se em Rio Branco (75%), alastrando-se pela capital vizinha Manaus (62,5%) e pelas mais distantes Porto Velho/RO (37,5%) e Boa Vista (RR) (37,5%), com a menor produtividade, no Norte, no Amapá, em sua capital Macapá (25%).

Assim, vê-se o desenhar de uma área dialetal formada pela distribuição espacial do item lexical *mucura* no Norte e no Nordeste do Brasil, o que demonstra que a dicionarização dessa marca de uso em Houaiss (2001) condiz com os dados aqui analisados, bem como nota-se também que os falantes comportam-se de maneira particular na nomeação dos elementos que os cercam, utilizando-se de designativos diversos para o mesmo referente, no caso com a utilização de cinco unidades léxicas distintas para nomear o “bicho que solta um cheiro ruim quando se sente ameaçado” (QSL/ALiB 71).

Outra unidade lexical cuja distribuição diatópica demonstra a presença de áreas dialetais é a *muriçoca*, utilizada para designar o “inseto de perninhas compridas que canta no ouvido das pessoas à noite” (QSL/ALiB 88). Sendo o segundo item lexical mais produtivo no Brasil como resposta para a pergunta 88 (QSL/ALiB) (27,6% do total de ocorrências), superada apenas por *pernilongo* (34,9%), *muriçoca* foi produtiva em todas as regiões brasileiras, à exceção do Sul, mas de maneira muito particular em cada área. A dimensão diatópica do item léxico em causa é demonstrada pela Carta Linguística VI, a seguir:

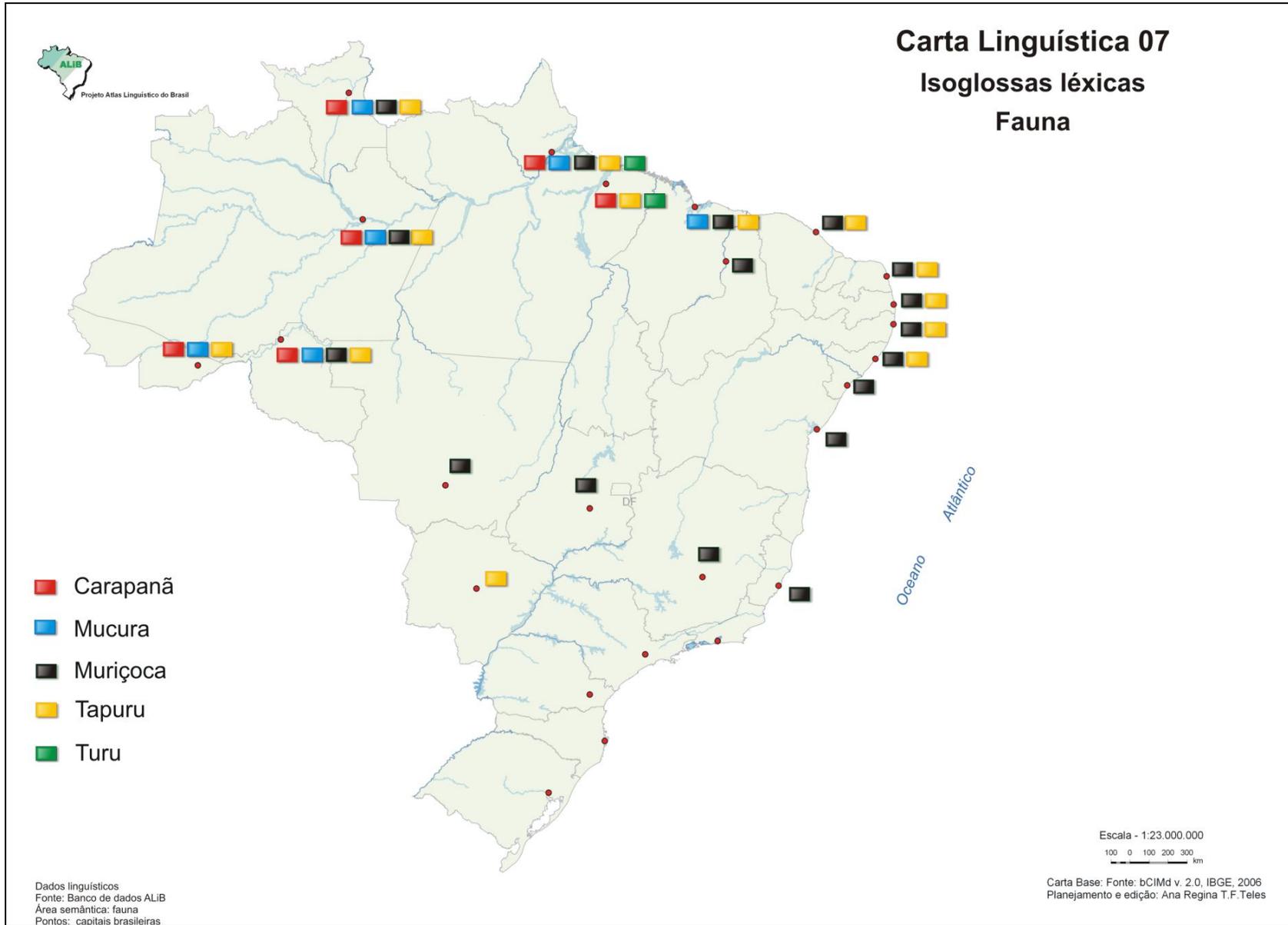


Observa-se, pela Carta apresentada, que *muriçoca* desenha-se num *continuum* geográfico, partindo do Nordeste para o Norte e norte da região Centro-Oeste, e a leste da região Sudeste. Essa disseminação do item léxico *muriçoca* por regiões para além da dicionarizada (Nordeste) pode ser atribuída às migrações de nordestinos para outras regiões do Brasil.

Além disso, verifica-se, também por meio da análise da Carta Linguística VI, que o léxico, de fato, permanece em constante mutabilidade. Prova disso é a maior ocorrência de *muriçoca* na capital goiana (80%), em detrimento inclusive das capitais nordestinas (nessa região, o maior percentual de ocorrência foi em João Pessoa/PB, com 66,6%). Isso sugere que outras designações estão adentrando no território em que antes predominava essa unidade lexical, no caso, o nordeste brasileiro.

Analisando-se, pois, as unidades léxicas de origem indígena documentadas na área semântica da *fauna* e sua distribuição diatópica, nota-se que puderam, de fato, ser identificadas isoglossas léxicas em território nacional, tal qual aventado como hipótese no delinear deste trabalho. Isso porque algumas delas tiveram seu uso concentrado em certas regiões do Brasil, demonstrando uma particularização no falar de determinadas localidades brasileiras, especialmente no Norte e Nordeste do Brasil.

Itens lexicais como *muriçoca* e *mucura*, já apresentados anteriormente, demonstraram essas isoglossas, bem como outros, como *carapanã*, *tapuru* e *туру*. Dessa maneira, para uma melhor visualização dessas linhas imaginárias, elaborou-se a Carta Linguística VII, apresentada na sequência:



Assim, identifica-se que *tapuru* e *muriçoca* disseminaram-se especialmente pelo Norte e Nordeste do Brasil, tendo a segunda unidade léxica se registrado também no Centro-Oeste e Sudeste do Brasil. *Carapanã* e *туру*, por sua vez, concentraram-se apenas no Norte do país, a primeira dominando todas as capitais nortistas e a segunda, documentando-se apenas no Macapá (AP) e em Belém (PA). Também *mucura* teve sua concentração na região amazônica, tendo apenas um registro na nordestina São Luís (MA).

Frente ao exposto, nota-se, a partir da análise da Carta Linguística VII e das demais ponderações apresentadas no decorrer da análise diatópica das unidades lexicais em causa, que as regiões Norte e Nordeste apresentam um falar caracterizado por certa particularização, no que diz respeito ao uso dos indigenismos aqui estudados. De fato, esses indigenismos apresentados concentraram-se nessas regiões, disseminando-se, por vezes, mas não havendo o registro de outras unidades lexicais que caracterizasse isoglossas léxicas nas demais regiões brasileiras.

De acordo com os estudos realizados, isso pode ser justificado, seja pela característica nordestina de conservadorismo linguístico e de ser essa região fonte de fluxos migratórios, seja pelo fato de o Norte caracterizar-se pela pouca mobilidade de seus habitantes e mesmo por agregar a maior população indígena brasileira, falante ainda nos dias atuais de línguas indígenas.

E finalizando esta etapa da análise dos dados, faz-se necessário assinalar, com base nos registros das unidades lexicais documentadas na área semântica da *fauna* e da sua dicionarização, no que diz respeito às marcas de uso, que o estudo ratificou a importância do léxico como o nível linguístico que melhor revela a dinamicidade da língua em uso, ao mesmo tempo em que

confirmou, por vezes, as marcas de uso documentadas nos dicionários gerais da língua portuguesa, mas também, em alguns casos, demonstrou que essa dicionarização fica aquém do real uso de certos designativos, como pode ser visto no Quadro XII a seguir:

| Item lexical | Houaiss (2001) | Ferreira (2004) | Dados ALiB (capitais brasileiras) |
|-------------------|--|--|-----------------------------------|
| Maritaca | Regionalismo do Sudeste | Brasileirismo da Amazônia | CO/Goiânia ⁵² |
| Ajuru | Regionalismo amazônico | Sem registro de marca dialetal | N/Macapá ⁵³ |
| Jaratataca | Regionalismo do Brasil | Sem registro de marca dialetal | CO/Campo Grande ⁵⁴ |
| Mucura | Regionalismo da Amazônia | Sem registro de marca dialetal | Norte ⁵⁵ |
| Timbu | Regionalismo do Brasil | Brasileirismo do Nordeste | NE/Recife, Natal e N/Porto Velho |
| Sarué | Regionalismo do Nordeste | Brasileirismo do Nordeste | SE/Vitória ⁵⁶ |
| Capenga | Regionalismo do Brasil | Sem registro de marca dialetal | N/NE/CO/SE/S |
| Tapuru | Regionalismo de Mato Grosso (cupim) | Brasileirismo de Mato Grosso (bicho da fruta) | N/NE |
| Turu | Não dicionarizado na acepção de bicho de pau podre | Não dicionarizado na acepção de bicho de pau podre | N/ Macapá e Belém |
| Muriçoca | Regionalismo do Nordeste e de Minas Gerais | Brasileirismo do Nordeste e de Minas Gerais | NE |
| Maruim | Regionalismo do Brasil | Sem registro de marca dialetal | NE/Maceió, São Luís e Natal |
| Carapanã | Regionalismo Amazônico | Brasileirismo Amazônico | Norte |

Quadro XIII – Marcas dialetais dicionarizadas e documentadas pelo Projeto ALiB - itens lexicais da área semântica da *fauna*

⁵² Ocorrência única.

⁵³ Houve apenas uma ocorrência.

⁵⁴ Ocorrência única. As unidades lexicais que tiveram baixa produtividade não foram consideradas como demonstrativas de marcas de uso, tendo em vista sua pouca utilização, por parte dos informantes.

⁵⁵ No Norte, apenas em Belém essa unidade lexical não foi documentada.

⁵⁶ Houve apenas uma presença.

Nota-se que os dados geolinguísticos têm muito a contribuir para a Lexicografia, o que pode ser confirmado pelos dados deste estudo que trouxe à baila tanto o uso de unidades lexicais já dicionarizadas com marcas de uso, quanto itens lexicais que apontam para a inovação lexical no português do Brasil, seja pela identificação de novas acepções atribuídas pelo uso a palavras já existentes, seja pela descoberta de neologismos que estão enriquecendo o vocabulário dos falantes das capitais brasileiras.

A área semântica da *fauna* mostrou-se assim muito produtiva, no que diz respeito ao registro de indigenismos, especialmente de origem tupi, fato comprovado pelo número de unidades lexicais identificadas: as 25 unidades léxicas estudadas são dessa base linguística.

Assim, concluindo-se a análise dos dados da área semântica da *fauna*, na sequência do trabalho, são analisadas as respostas para as perguntas do QSL/ALiB vinculadas à área semântica da *alimentação e cozinha*.

4.3 ANÁLISE DE DADOS DA ÁREA SEMÂNTICA ALIMENTAÇÃO E COZINHA

Esse recorte é composto por 12 questões que foram respondidas pelos 200 informantes entrevistados nas 25 capitais brasileiras que fazem parte da rede de pontos do Projeto ALiB. Dentre as doze questões, em cinco delas foram documentadas unidades léxicas de base linguística indígena.

| Número da questão | Pergunta | Indigenismo documentado |
|-------------------|--|---|
| 179 | “Como se chama uma papa cremosa feita com coco e milho verde ralado, polvilhada com canela?” | <i>Curau</i> <i>mingau</i> <i>mingau de milho</i> <i>mingau de milho verde</i> <i>pamonha</i> |
| 180 | “E essa mesma papa, com milho verde ralado, sem coco, como é que chama?” | <i>curau/curau de milho</i> <i>mingau</i> |

| | | |
|-----|---|---|
| | | <i>mingau de milho</i> <i>pamonha</i> |
| 181 | “Como se chama aquele alimento feito com grãos de milho branco, coco e canela?” | <i>Mingau</i> <i>mingau de milho</i> <i>mingau de milho branco</i> <i>mingau de mungunzá</i> <i>mingau maranhense</i> |
| 182 | “Como se chama a bebida alcóolica feita de cana-de-açúcar?” | <i>caxiri</i> <i>pitu</i> <i>jacaré</i> <i>uca</i> |
| 186 | “Como se chama isto? (pão francês)” | <i>carioca/carioquinha</i> |

Quadro XIV: Indigenismos documentados na área semântica *alimentação e cozinha* nas capitais brasileiras

Porém, antes de serem analisados os indigenismos registrados, objeto deste trabalho, faz-se necessário localizá-los, em termos de produtividade, em relação ao cômputo geral das respostas informadas. Nesse sentido, foram documentadas, no total, 188 unidades léxicas para designar os conceitos em questão. Dessa forma, os 08 indigenismos – *carioca/carioquinha*, *caxiri*, *curau*, *jacaré*, *mingau/variantes*, *pamonha*, *pitu* e *uca* – correspondem a 4,3% dos itens lexicais documentados. A distribuição de todas as unidades léxicas, de acordo com sua base linguística, é visualizada no Gráfico VII, a seguir:

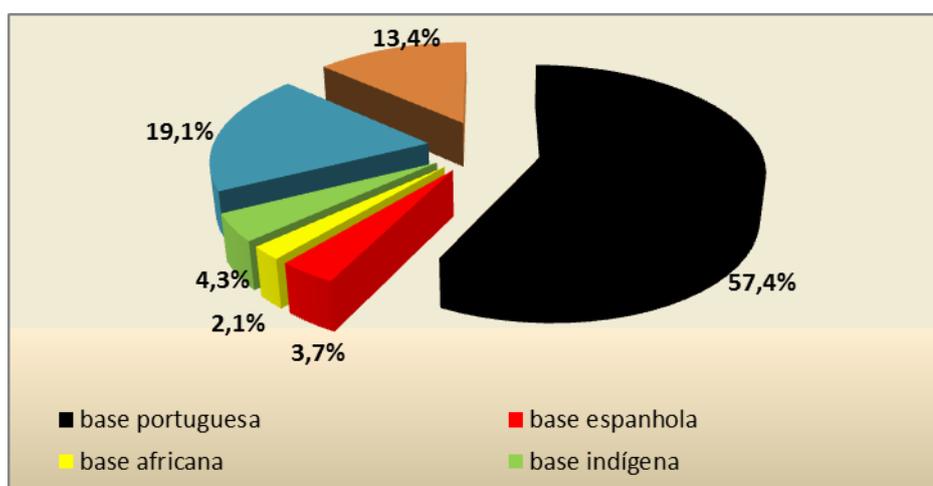


Gráfico VII: Distribuição das unidades lexicais relacionadas à área semântica da *alimentação e cozinha* de acordo com a base linguística

Para fins de cálculo da produtividade, as variantes *carioca* e *carioquinha* foram associadas de modo a representarem apenas uma unidade léxica. Assim também ocorreu com as variações de *mingau*, conforme especificadas durante a análise semântica, tendo em vista o fator proximidade/distanciamento do referente em questão. A exemplo do adotado no exame dos dados das demais áreas semânticas, a discussão do assunto respeitou a ordem das perguntas que deram origem aos dados.

Nesse sentido, a primeira unidade lexical a ser analisada sob o viés léxico-semântico é *mingau*. Segundo Cunha (2010), esse designativo de base tupi advém do étimo **mina’u** e significa “alimento de consistência pastosa, espécie de papa preparada com farinha de mandioca ou de trigo (ou fubá, maisena, aveia etc.), diluída e cozida em água ou em leite e a que se adicionam açúcar, ovos, canela, etc.”. Sampaio (1928), por sua vez, atribui-lhe o étimo tupi **mingá-ú**, que nomeia “o comer visguento, as papas ralas de mandioca” (SAMPAIO, 1928). Houaiss (2001), por seu turno, acrescenta-lhe, por extensão de sentido, a acepção “qualquer substância com consistência de mingau”.

Considerando-se, pois, que se busca o nome para a “papa cremosa feita com coco e milho verde ralado, polvilhada com canela”, nota-se que a definição de *mingau* apresentada por Houaiss (2001) aproxima a acepção ao perguntado pelas questões 179 e 180, esta última que busca designativos para “essa mesma papa, com milho verde ralado, sem coco”, uma vez que trata da consistência do alimento, já que, em Cunha (2010), tem-se que os ingredientes-base da receita são diversos do milho⁵⁷. Já a pergunta 181 – “como se chama

⁵⁷ Cunha (2010) registra como ingredientes “farinha de mandioca ou de trigo (ou fubá, ou maisena, aveia etc.)”.

aquele alimento feito com grãos de milho branco, coco e canela?” – que também motivou a ocorrência de *mingau*, busca nomes para um referente que não é necessariamente uma papa, um mingau.

Talvez também por esse motivo *mingau* venha associado, em parte, a elementos especificadores, como *mingau de milho*, *mingau de milho verde* e *mingau de milho branco*, evidenciando o ingrediente milho na preparação. Já *mingau de mungunzá* traz, como segundo elemento, um dos designativos buscados pela QSL/ALiB 181.

O item lexical *mungunzá* não foi analisado neste trabalho por ser uma unidade lexical de base africana. Trata-se de um dos nomes para a “espécie de mingau feito de milho branco com leite e leite de coco, temperado com açúcar e canela”. Nesse sentido, *mingau de mungunzá* seria, de acordo com a análise semântico-lexical, próximo do referente descrito pela questão 181 (QSL/ALiB).

Também *mingau maranhense* nomeia o mesmo doce designado no Norte e no Nordeste como *mungunzá*. Pesquisando a culinária do Maranhão, verificou-se que o alimento assim designado é preparado com milho branco, açúcar, leite condensado e polvilhado com canela⁵⁸. Contudo, esse item lexical não foi mencionado pelos maranhenses, mas por dois informantes de Teresina (Piauí), o que pode demonstrar o conhecimento dos hábitos da região pelos habitantes do Estado vizinho, ou mesmo que um nome ultrapassou as fronteiras geográficas e passou a ser conhecido em outras localidades, mesmo não sendo mais frequente no local de origem.

⁵⁸ Informação obtida por meio de consulta aos sites de culinária maranhense <http://cozinhamaranhense.blogspot.com.br/2010/04/mingau-de-milho-maranhense.html> e <http://www.almanaqueculinario.com.br/receita/doces-e-sobremesas/mingau-de-milho-maranhense-8945.html>. Acesso em: 14 maio 2012.

Assim, vê-se que os informantes fizeram uso de um nome genérico para nomear os referentes em questão, uma vez que *mingau* foi produtivo como resposta para as perguntas 179, 180 e 181 do Projeto ALiB. Já *mingau de milho* apareceu apenas em duas dessas, 179 e 180, já especificando o ingrediente principal do prato. Em *mingau de milho verde* (QSL/ALiB 179), *mingau de milho branco* (QSL/ALiB 180) e *mingau de munguzá*, por sua vez, ocorre a especificação do milho de que é feito o prato.

Quanto à unidade lexical *curau*, a próxima a ser analisada, Machado (1987), em seu *Dicionário etimológico da Língua Portuguesa*, atribui-lhe a origem tupi, sem, todavia, apontar o étimo da palavra. As demais obras etimológicas consultadas não registram essa unidade lexical. Em relação à dicionarização, Houaiss (2001) não contempla essa unidade léxica na sua nomenclatura, mas no verbete *canjica*, termo definido como “papa cremosa de milho verde ralado e cozido com leite e açúcar” (HOUAISS, 2001), lista-a como sinônimo.

A unidade lexical *pamonha*, por seu turno, também foi utilizada para nomear os referidos alimentos, porém, não sendo resposta válida para a pergunta que motivou o seu uso, uma vez que nomeia a “espécie de bolo feito de milho verde, leite de coco, manteiga, canela, erva-doce e açúcar, e cozido em tubos das folhas do próprio milho ou de folhas de bananeira, atados nas extremidades” (FERREIRA, 2004) e não o alimento em causa, não foi aqui analisada, apenas citada em sua base etnolinguística, tendo em vista que se trata de tupinismo de étimo **pamu’ñã**, segundo Houaiss (2001).

Já *caxiri*, *pitu*, *jacaré* e *uca* são variantes léxicas utilizadas para nomear “a bebida alcóolica feita de cana-de-açúcar”. *Caxiri* é apontada por Cunha (1989) como de provável origem tupi, sem, contudo, apresentar seu étimo definido,

informação também não encontrada em Sampaio (1928), que não documenta a unidade lexical em questão. Segundo Ferreira (2004), *caxiri* designa a própria cachaça, informação partilhada também por Houaiss (2001).

Pitu, por sua vez, também de origem tupi, e cujo étimo, **pi'tu**, segundo Cunha (2010), nomeia uma espécie de camarão, foi provavelmente utilizada para designar a aguardente tendo em vista que se trata de uma marca da bebida alcoólica à venda no mercado, razão pela qual, por um processo metonímico, o nome próprio designou o nome comum da bebida, isto é, foi utilizada a marca para nomear o produto.

Também esse é o caso de *jacaré*, resposta para a pergunta 182 do QSL/ALiB. Do étimo tupi **ya-caré**, “aquelle que é orto, ou sinuoso” (SAMPAIO, 1928), nomeia o réptil da família dos “crocodilianos aligatorídeos” (FERREIRA, 2004) e é marca de aguardente, tendo sido mencionada por uma idosa soteropolitana.

Já *uca*, do étimo **u'sa** (HOUAISS, 2001), além da acepção de “caranguejo”, consta no dicionário de Ferreira (2004) como gíria para cachaça e, na segunda entrada de Houaiss (2001), vigora como “aguardente de cana, cachaça”. Enriquecendo essa vertente de sentido, encontra-se, no Dicionário Online de Língua Portuguesa (Disponível em: <http://www.dicio.com.br/uca/>), *uca* como sinônimo de cachaça, assim como no site Cultura e Conhecimento (COSTA, 2012), no artigo História da Cachaça, em que vigora como a mesma bebida, dentre outros sinônimos.

A última unidade lexical da área semântica *alimentação e cozinha* a ser analisada sob a ótica léxico-semântica é *carioca/carioquinha*. Utilizada para nomear o pão comumente conhecido como pão francês, *carioca*, como já

mencionado no tópico acerca das *atividades agropastoris*, que também registrou essa unidade léxica para nomear a tangerina, advém da palavra tupi **kari'oka**, que significa casa do branco (CUNHA, 2010), e diz respeito aos indivíduos nascidos no Estado do Rio de Janeiro. É importante citar também que essa unidade lexical foi mencionada por um morador de Manaus (Amazonas) e, especialmente, por todos os informantes de Fortaleza (Ceará), demonstrando tratar-se de um item da norma lexical dos habitantes da capital cearense para nomear o pão também conhecido como *francês*.

Uma provável motivação para o nome *pão carioca* pode advir da origem brasileira do conhecido pão francês. O alimento é diferente do pão da França, por possuir açúcar e gordura como ingredientes, e foi produzido inicialmente no Rio de Janeiro a pedido de afortunados brasileiros que voltavam da Europa por volta do período da Primeira Guerra Mundial, na década de 1920⁵⁹.

Encerrando a análise semântico-lexical do *corpus* relacionado à área semântica da *alimentação e cozinha* aqui estudado, algumas considerações podem ser tecidas, em caráter de síntese. A primeira delas é que, como Barbosa ([s.d.] *apud* HOUAISS, 1985) já observara, a presença de indigenismos de origem não tupi é bem menos expressiva na variante brasileira do português que os tupinismos.

Endossa essa constatação o recorte aqui estudado, à medida que revelou que os oito indigenismos estudados foram de base linguística tupi, demonstrando que a base da Língua Geral Paulista, que por muito tempo foi a língua mais falada no país, ainda está expressa nas escolhas lexicais dos falantes das 25 capitais brasileiras entrevistadas, dando-lhe cor local.

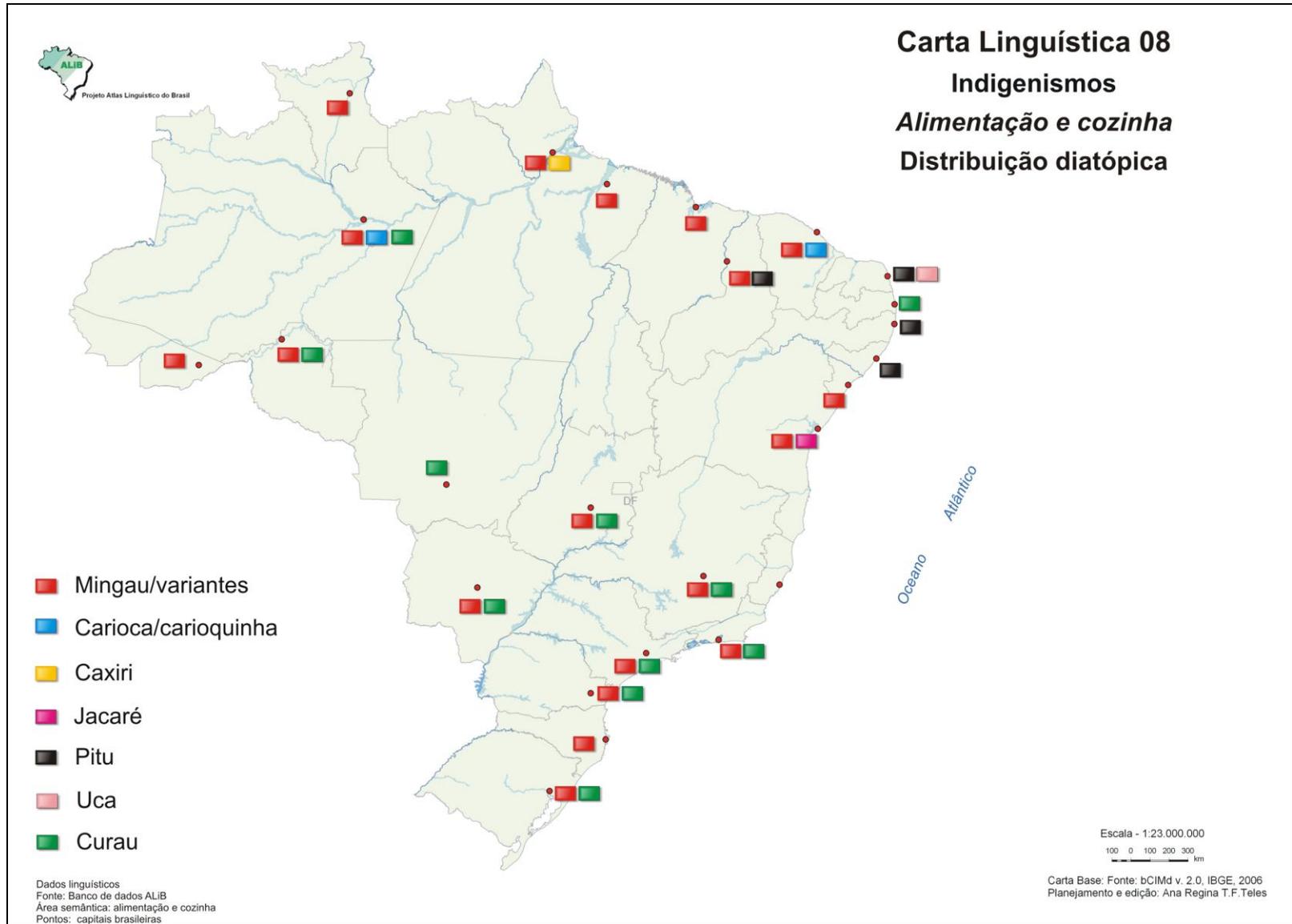
⁵⁹ Segundo a enciclopédia virtual Wikipédia, disponível em: http://pt.wikipedia.org/wiki/P%C3%A3o_franc%C3%AAs.

Além disso, percebeu-se a escolha, por parte dos falantes, de unidades lexicais por vezes mais genéricas para nomear os referentes em pauta, como o uso de *mingau*, mas também, em algumas situações, acompanhadas de elementos especificadores, como *mingau de milho branco* e *mingau de mungunzá*, por exemplo.

Realizadas essas considerações, a continuidade do trabalho se dará com a análise diatópica das unidades léxicas em questão. Para isso, será necessário considerar que, após essa análise primeira, *pamonha* foi excluída da análise diatópica, o que resulta num total de 07 itens lexicais a vigorarem na análise a seguir.

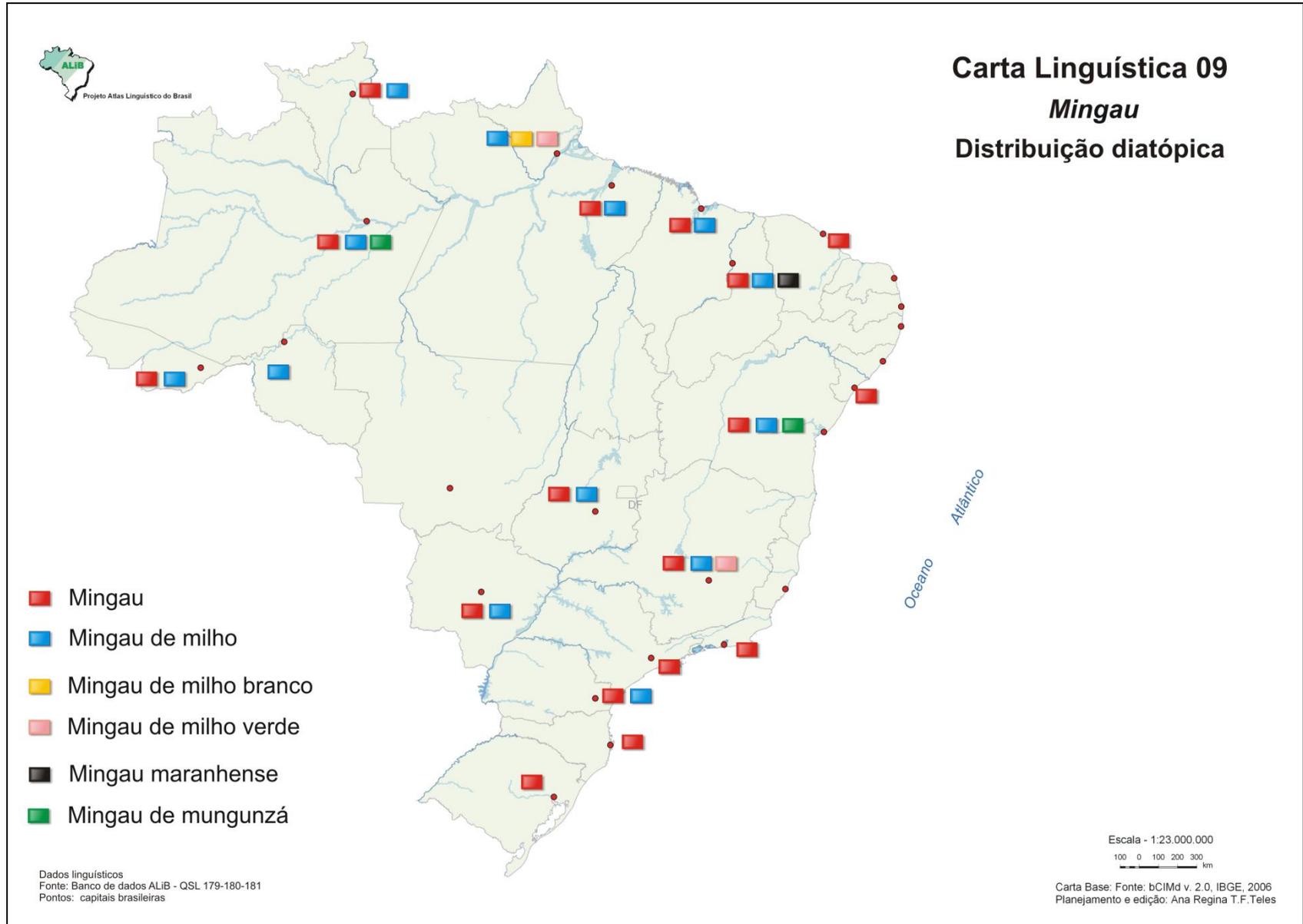
Para visualizar a distribuição espacial desses 07 indigenismos, elaborou-se a Carta Linguística VIII, apresentada na sequência⁶⁰.

⁶⁰ Na legenda, a unidade lexical *mingau* foi associada às demais variantes formadas com o item lexical *mingau*.



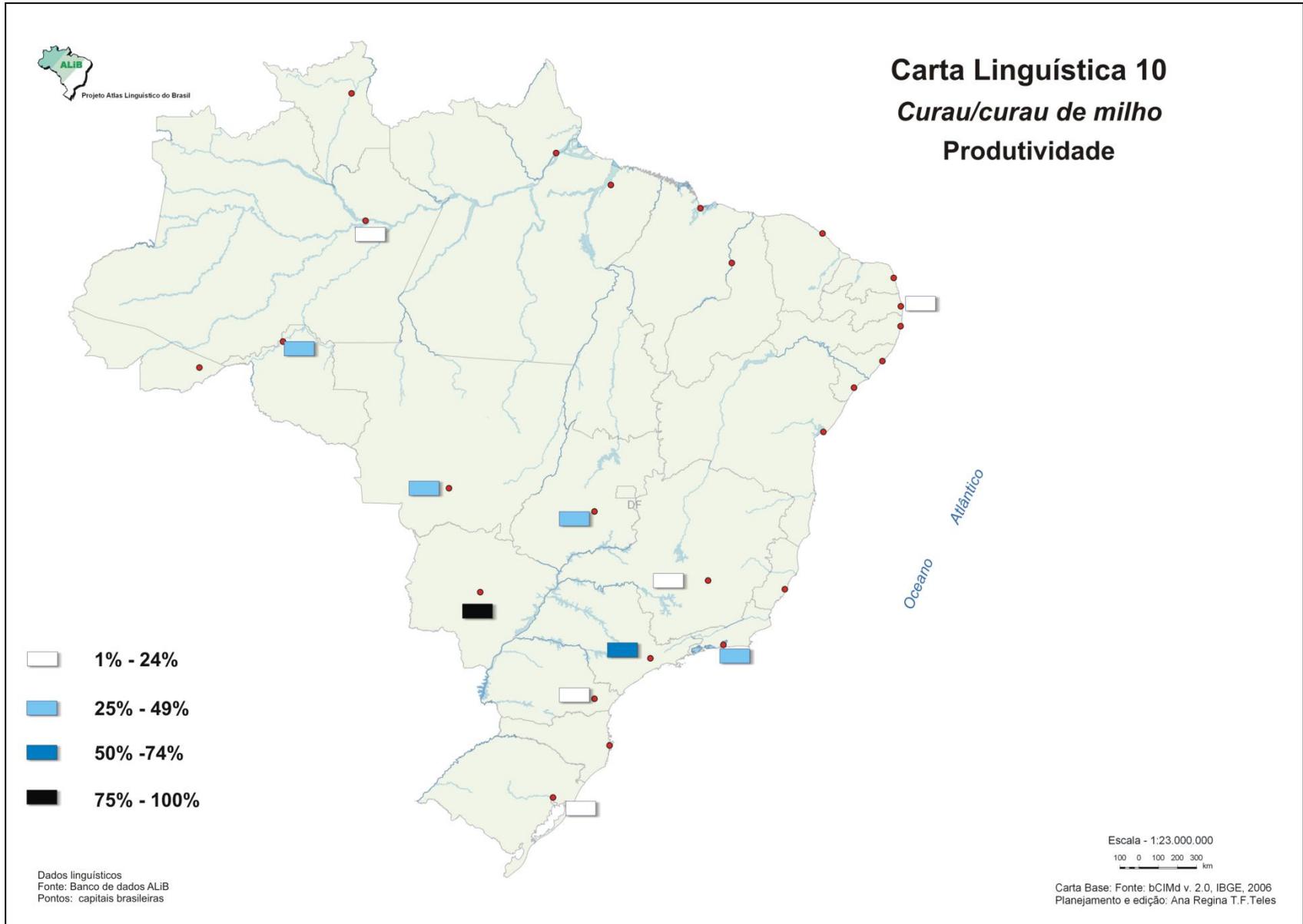
Observando-se a Carta Linguística VIII, vê-se que o item lexical genérico *mingau* usado para nomear o alimento descrito pelas perguntas 179 (papa cremosa feita com coco e milho verde ralado, polvilhada com canela), 180 (a mesma papa, com milho verde ralado, sem coco) e 181 (alimento feito com grãos de milho branco, coco e canela) está disseminado por todas as regiões brasileiras.

Ao dissociarmos as variantes de *mingau* nas seis categorias em que foram documentadas – *mingau*, *mingau de milho*, *mingau de milho branco*, *mingau de milho verde*, *mingau maranhense* e *mingau de mungunzá* –, notamos que o item lexical *mingau* ocorreu num maior número de capitais (17), seguida de perto por *mingau de milho* (13). Já *mingau de milho verde* e *mingau de mungunzá* foram registradas em 02 capitais, cada um, enquanto *mingau de milho branco* foi mencionado apenas em Macapá e *mingau maranhense* somente em Teresina. A distribuição espacial dessas ocorrências está melhor visualizada na Carta Linguística IX, na sequência:



O item lexical *curau*, por sua vez, foi utilizado para nomear os conceitos expressos pelas questões 179 e 180, tem sua distribuição concentrada no Centro-Sudeste do Brasil: foi mencionado, na região Nordeste, por apenas um informante, da cidade de João Pessoa (PB); no Norte, foi citado por um manauara e mais dois portovelhenses. Já na região central do Brasil, foi mencionado nas três capitais do Centro-Oeste (25% de ocorrência em Cuiabá, 37,5% em Goiânia e 75% em Campo Grande). Em relação ao Sudeste, 04 dos 08 informantes paulistas proferiram essa resposta (50%), enquanto no Rio de Janeiro 02 falantes a mencionaram (25%), e em Minas Gerais houve apenas uma menção. No Sul, por seu turno, apenas um curitibano e um porto-alegrense utilizaram *curau* para nomear a papa cremosa, com e sem coco.

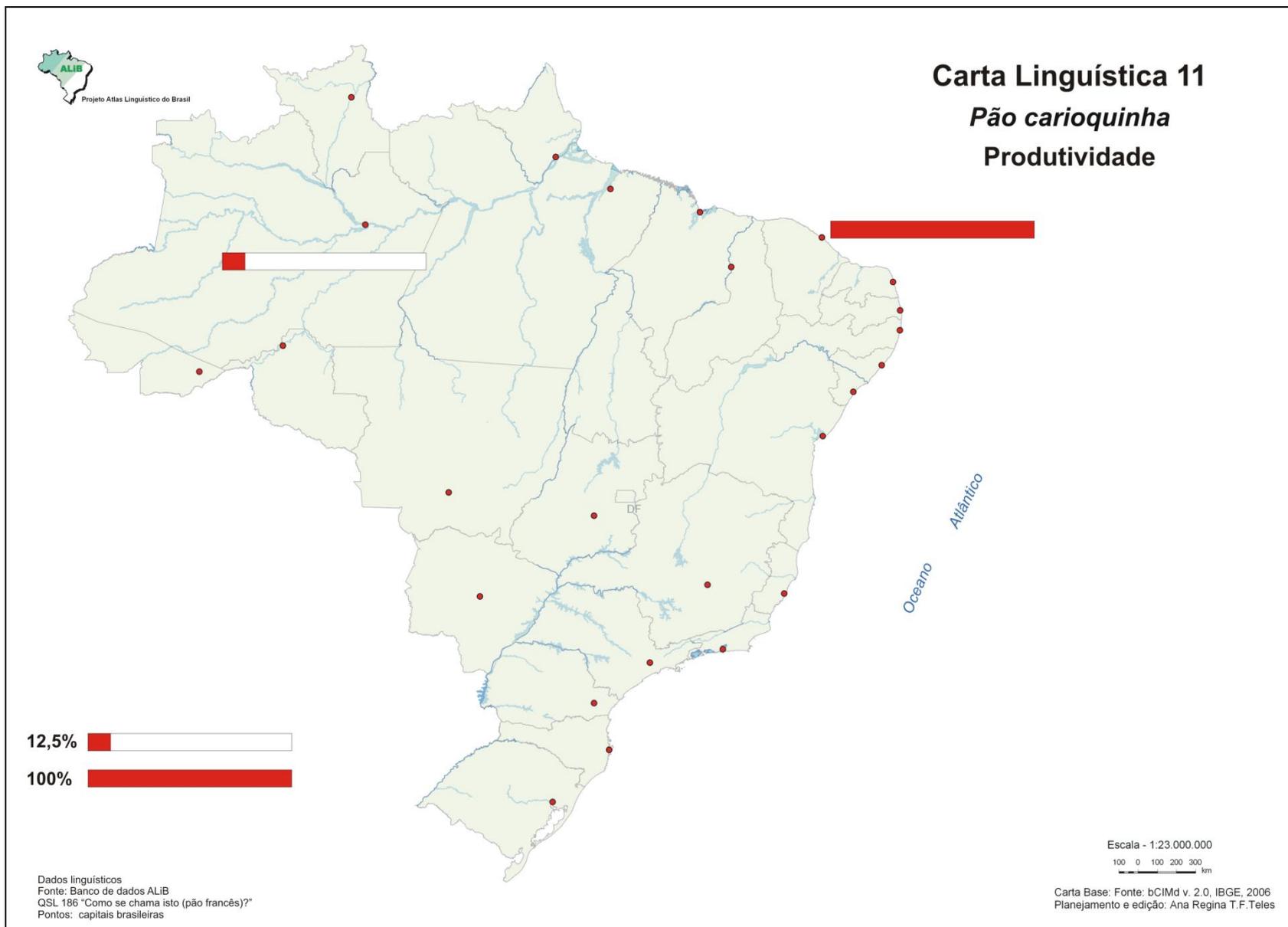
Assim, nota-se que o uso de *curau* está concentrado na região central do Brasil, desenhando um *continuum* de irradiação de incidência entre Campo Grande (MS) e São Paulo (SP), e se alastrando e diminuindo sua produtividade à medida que se afasta dessas localidades, além disso, teve uma ocorrência em João Pessoa (PB). Esse *continuum* geográfico pode ser melhor visualizado na Carta Linguística X, a seguir:



Já os demais itens lexicais estão concentrados em áreas específicas. A unidade lexical *caxiri*, por exemplo, foi utilizada para nomear a bebida comumente conhecida como cachaça pelo informante jovem de alta escolaridade em Boa Vista, enquanto o item lexical *jacaré* foi citado por uma mulher, soteropolitana, idosa, de baixa escolaridade. O item lexical *pitu*, por sua vez, foi mencionado pelo maceioense idoso de baixa escolaridade e pelos jovens de alta escolaridade do Recife, além das idosas de baixa escolaridade recifense e teresinense. Além disso, dois natalenses nomearam a cachaça como *pitu*, o jovem e a idosa de baixa escolaridade.

Além de designarem o mesmo referente, outra característica partilhada por esses três itens lexicais é que foram utilizados como segunda e terceira respostas⁶¹ (*caxiri*, *jacaré* – segunda resposta, e *pitu* – terceira resposta), sempre antecedidas pelo item lexical *cachaça*, que parece ser o termo de uso categórico para designar a bebida em pauta, não analisado no âmbito deste trabalho por ser de base portuguesa. Já as unidades lexicais *pão carioca* e *carioquinha* são tratados a partir da Carta Linguística XI, visualizada a seguir:

⁶¹ Em relação à ordem das respostas, é importante explicar que, quando da aplicação do questionário, os pesquisadores do Projeto ALiB buscam o maior número de unidades léxicas para nomear os conceitos requisitados pelas perguntas. Assim, quando o informante responde a pergunta, comumente completa-se a questão com “conhece outro nome para...”; dessa forma, investigam-se todas as possíveis designações utilizadas pelos falantes, e a ordem na menção pode revelar maior/menor uso dos itens lexicais informados.



Observa-se que a unidade lexical *pão carioca*, apresentada na Carta Linguística XI, teve produtividade altamente relevante em Fortaleza, capital do Estado do Ceará. Nessa localidade, todos os informantes mencionaram *pão carioca/carioquinha* como resposta quando perguntados sobre o nome do pão conhecido também como francês. Inclusive, a unidade léxica em pauta foi mencionada como primeira resposta por sete dos oito informantes entrevistados, confirmando seu uso naquela cidade. Já em Manaus/AM, o uso desse item lexical pode evidenciar influência dos contatos entre diferentes normas regionais, tendo em vista que muitos nordestinos migraram para o Norte, sendo provável o contato destes com a jovem jornalista entrevistada pela equipe do Projeto ALiB em Manaus.

Comparando-se as definições dicionarizadas e a distribuição espacial dos itens lexicais, segundo o banco de dados do Projeto ALiB, nota-se que um dos objetivos deste trabalho, que é o de contribuir com os trabalhos lexicográficos no que diz respeito às marcas de uso dessas unidades léxicas, é atendido à medida que a área semântica da *alimentação e cozinha* muito tem a acrescentar ao universo já dicionarizado dessas marcas de uso. Isso porque se viu que *mingau*, não listado nos dicionários como o nome dos alimentos preparados à base de milho (179, 180 e 181 QSL/ALiB), teve presença disseminada por todo o Brasil, seja como forma simples, seja por meio das formas compostas *mingau de milho, de milho verde, de milho branco, maranhense* e *de mungunzá*. Os itens lexicais *carioca/carioquinha* também, apesar de não estarem registrados em nenhuma das obras lexicográficas consultadas na acepção de pão francês, mostraram-se produtivos para designar esse referente em Fortaleza (CE).

Já *pitu*, que nas obras consultadas só consta como um tipo de camarão (HOUAISS, 2001), teve registro considerável no Nordeste brasileiro para nomear a bebida comumente chamada de cachaça (03 informantes do Recife (PE), 02 em Natal (RN) e Maceió (AL) e mais 01 em Teresina (PI) mencionaram tal item léxico). O Quadro XIV organiza todas essas informações, sintetizando a dicionarização e as características regionais reveladas pelos dados ora analisados com vistas a contribuir para trabalhos lexicográficos, no que tange ao registro de marcas de uso.

| Item lexical | Houaiss (2001) | Ferreira (2004) | Dados ALiB (capitais brasileiras) |
|-----------------------------------|---|---|--|
| <i>Carioca/carioquinha</i> | Não dicionarizado na acepção de pãozinho de água e sal. | Não dicionarizado na acepção de pãozinho de água e sal. | NE/Fortaleza (pãozinho de água e sal) |
| <i>Curau</i> | Canjica (papa) | Canjica (papa de consistência cremosa feita com milho verde ralado, a que se acrescenta açúcar, leite de vaca ou de coco, e polvilha com canela). | N/NE/CO/SE/S |
| <i>Pitu</i> | Não dicionarizado na acepção de bebida feita de cana de açúcar. | Não dicionarizado na acepção de bebida feita de cana de açúcar. | NE (bebida feita de cana de açúcar) |

Quadro XV – Marcas dialetais dicionarizadas e documentadas pelo Projeto ALiB - itens lexicais da área semântica da *alimentação e cozinha*

Nota-se assim que, como já sabido, a língua comporta-se como um organismo vivo, em constante atividade, revelando traços por vezes ainda não registrados em obras, como é o caso da identificação das marcas de uso para as unidades lexicais em causa nos dicionários pesquisados, aquém do que os dados do Projeto Atlas Linguístico do Brasil têm revelado. Isso demonstra também que

os trabalhos geolinguísticos em muito têm a auxiliar os estudos lexicográficos, demonstrando a importância dessas pesquisas para o conhecimento e registro da língua em uso num dado espaço e tempo.

Por fim, vê-se, a partir dessas considerações e comparando-se os dados relativos à área semântica *alimentação e cozinha* com os relacionados à *fauna* e às *atividades agropastoris*, nota-se uma maior produtividade de indigenismos nestas áreas semânticas frente àquela, tendo que se considerar também que essas áreas semânticas possuem mais questões que *alimentação e cozinha*. Todavia, no que diz respeito à origem das unidades lexicais examinadas, observa-se que, em todas as áreas, a totalidade é de base tupi, o que demonstra que essa etnia deixou traços marcantes e perenes na cultura e na norma lexical brasileira, o que se pode observar a partir do recorte analisado no âmbito deste trabalho.

CONSIDERAÇÕES FINAIS

A língua, patrimônio cultural de um grupo social, sempre traz arraigados consigo fatores que perpassam os limites linguísticos em sua composição. O processo de nomeação da realidade revela, pois, também a forma de pensar do homem, uma vez que, ao nomear o mundo que o cerca, ele imprime, nos designativos, seus conhecimentos, crenças e tradições, sua cultura, enfim, formando seu repertório lexical.

E, no Brasil, as etnias que participaram ativamente da formação da cor local da população brasileira deixaram marcas no léxico em uso nas capitais brasileiras, representadas pela presença de indigenismos na norma lexical ativa dos habitantes dessas localidades. O número de unidades léxicas de base indígena documentado a partir do *corpus* aqui analisado demonstrou que o processo de nomeação, a partir do qual o homem estrutura sua realidade, está diretamente influenciado pela presença indígena na nação brasileira, desde o Período Colonial até os dias atuais.

Assim, a complexidade cultural refletida pelo léxico, conforme assevera Sapir (1969), é assim também verificada na norma lexical brasileira, considerando-se que “a diversidade étnica dos grupos populacionais brasileiros [...] acabou provocando o exercício de trocas ou empréstimos vocabulares, para nos termos, aqui, apenas ao campo léxico” (DICK, 2000, p. 305), realidade essa confirmada pelas 58 unidades lexicais de base indígena aqui analisadas que são

representativas da variante brasileira do português utilizado pelos habitantes das capitais brasileiras.

Este trabalho evidenciou que os tupinismos são as únicas unidades léxicas de origem indígena representativas nos dados analisados, representando 100% das ocorrências de indigenismos aqui estudados. Assim, vê-se que as línguas indígenas diversas do tupi estão pouco representadas no Português do Brasil, pelo menos no que diz respeito aos dados aqui analisados, coletados pelo Projeto Atlas Linguístico do Brasil referentes a elementos da *fauna*, da *alimentação e cozinha* e das *atividades agropastoris*.

Sabe-se, retomando o conteúdo apresentado no Capítulo I deste trabalho, que os portugueses que aqui chegaram utilizaram os conhecimentos indígenas no trato com os animais e com a terra, nomeando os referentes próprios da nova terra, dentre eles os ligados às áreas semânticas aqui analisadas. Esse foi o motivo, inclusive, de se ter eleito o recorte do léxico estudado como objeto de investigação.

Além disso, e percorrendo os capítulos deste trabalho, as ciências que o embasaram, apresentadas no Capítulo II, foram extremamente relevantes para se chegar aos resultados atingidos, tendo em vista que o léxico reflete em alto grau os condicionantes sócio-históricos e geográficos. Isso porque, se a Lexicologia forneceu os subsídios para o estudo do léxico em si, a Geolinguística e a Dialetoлогия demonstraram esse aspecto social da língua, uma vez que

[...] A preocupação diatópica, seja porque os homens se situam, inevitavelmente, nos espaços geofísicos, seja porque as línguas e as suas variedades, pela implicações culturais a que estão sujeitas e que indubitavelmente as reflete, têm um território próprio, ou seja, ainda, porque o homem é indissociável no seu **existir** e no seu **agir**, no seu ser e no seu **fazer**, tem sido uma constante nos estudos dialetais e desde os seus primórdios (CARDOSO, 2010, p. 48).

Outra conclusão a que se pode chegar é que o percurso metodológico adotado, e explanado no Capítulo III desta dissertação, demonstrou que os estudos geolinguísticos têm sua importância confirmada ao documentar a língua em uso, como que registrando a fotografia da língua num dado espaço e momento histórico. Segundo Cardoso (2010, p. 169), essa importância reside no fato “de o conhecimento sistemático da variação, a delimitação de áreas linguísticas específicas e a relação entre os diferenciados usos da língua [...] constituírem-se num benefício de cunho social”, especialmente.

Quanto aos objetivos delimitados para a pesquisa que deram origem a este trabalho, nota-se que foram atingidos, a começar pelo objetivo geral, haja vista que se verificou a manutenção de indigenismos na norma lexical brasileira, aqui representada pelos dados orais coletados pelo Projeto ALiB referentes às áreas semânticas selecionadas para o estudo.

Também se pode comprovar que fatores como povoamento, identidade étnica e condicionantes geográficos interferem na norma lexical. Demonstra isso o fato de as unidades lexicais de base indígenas se concentrarem de forma mais representativa em locais em que o índio se destaca na constituição da população, como no Norte do país, região em que se documentaram 39 indigenismos, dos 58 estudados pela pesquisa.

Além disso, também a disseminação de unidades lexicais, por processos migratórios, por vezes, mostrou-se preponderante no uso de designativos indígenas para os elementos contemplados pelas perguntas do QSL selecionadas, como o uso de *muriçoca* se disseminando do Nordeste para as demais regiões, ou de *curau*, partindo do Centro-Oeste e subindo para o Norte e

Nordeste e descendo para o Sul do país, especialmente para as capitais próximas à região de sua maior ocorrência, Campo Grande, em Mato Grosso do Sul.

A disseminação das unidades lexicais demonstrou também a presença de isoglossas léxicas em território nacional, mostrando o caminhar das unidades léxicas pelo Brasil e sua maior ou menor concentração, de acordo com as localidades pesquisadas. Podem-se citar, além de *muriçoca* (área semântica da *fauna*) e *curau* (área semântica da *alimentação e cozinha*), também *mangará*, *mandioca*, *macaxeira* e *aipim* (*atividades agropastoris*) para comprovar a existência dessas linhas imaginárias que delineiam o falar do homem citadino entrevistado pelo Projeto Atlas Linguístico do Brasil.

E o estudo desse caminhar do léxico, revelado por este trabalho a partir de dados geolinguísticos, também visou a poder contribuir para os trabalhos lexicográficos, sobretudo no que diz respeito às marcas de uso dicionarizadas nas obras que trabalham com o léxico geral da língua. Os exemplos de unidades léxicas dicionarizadas como características de uma região, mas que foram documentadas em outras, demonstram que o repertório vocabular em uso permanece em constantes transformação e movimento, como revelam, para se citarem apenas alguns desses casos, o registro de *mucura* e *carapanã* – que confirmaram a marca de uso a elas atribuída pelos dicionaristas, e também *tapuru* – que se mostrou em área diversa da dicionarizada – e *mangará* – que se disseminou para área além da registrada nos dicionários.

E, além da contribuição para a documentação das marcas de uso, os dados analisados também podem contribuir para a própria dicionarização dos itens lexicais em pauta, pois revelaram novas acepções atribuídas a certas unidades lexicais de base indígena em uso pelos falantes das capitais brasileiras,

como o registro de *soca*, em Florianópolis, Santa Catarina, para nomear o sabugo de milho, e também *carioca* em Fortaleza, no Ceará, para designar o pão também conhecido como francês, por exemplo.

Assim sendo, verifica-se que o acervo vocabular permanece em constante adequação para o uso de seus falantes, que por vezes, como revelado pelos dados, atribuem novas acepções a unidades léxicas já existentes, assim como também demonstram o desuso de certas unidades lexicais, como a documentação de *jacá* principalmente na fala de idosos e preterida pelos jovens.

Essa constatação vai de encontro às considerações de Lucchesi (1998, p. 95), para quem “as mudanças na língua seriam, em sua maioria, uma resposta às novas exigências de expressão, decorrentes das alterações no universo cultural da comunidade linguística, das descobertas científicas e tecnológicas etc”. Essas alterações mencionadas pelo linguista podem também decorrer do distanciamento do falante de determinados referentes mais relacionados ao universo rural, como é o caso, dos contemplados pelas perguntas do Questionário Semântico-Lexical do Projeto ALiB, já que o desconhecimento do referente também pode desencadear a não resposta, em especial, entre os mais jovens, bem como podem gerar o considerável número de respostas não válidas tal qual se verificou neste estudo (04 na área semântica das *atividades agropastoris*, 05 na *fauna* e 01 na *alimentação e cozinha*).

Dessa forma, pode-se concluir que os 59 indigenismos analisados ao longo deste trabalho têm pouca participação no universo geral dos itens lexicais (7,2%); todavia, a pesquisa aprofundada que se pode fazer a partir deles revelou esses e outros traços marcantes em seu uso, como o aqui explanado, bem como

todas as considerações realizadas no Capítulo IV deste trabalho, que apresentou a análise dos dados apurados.

Porém, apesar desses resultados, devem-se também apontar os percalços enfrentados para que se pudesse concluir este trabalho. Além da dificuldade em se obterem obras que versassem sobre as línguas indígenas, desde sua história e classificação até seu registro, também os estudos etimológicos, por vezes, mostraram-se discrepantes quanto à etimologia das unidades lexicais. Casos como um mesmo item lexical estar assinalado como de duas ou mais origens distintas nas diversas obras estudadas, situações em que dicionários de língua revelavam uma origem indígena não confirmada e, além disso, unidades lexicais, cuja base linguística não foi encontrada em qualquer das obras pesquisadas (neste trabalho, foram 107 casos, representando 13,5% do total das respostas).

Isso demonstra a necessidade de maiores estudos, tanto das línguas indígenas como da etimologia do vocabulário da língua portuguesa em uso pelos brasileiros, aqui representados pelos 200 informantes das capitais brasileiras entrevistados pelo Projeto Atlas Linguístico do Brasil. E, nesses casos, como metodologia de trabalho adotada e apontada no Capítulo III desta Dissertação, usou-se o recurso de se considerarem os dicionários etimológicos que foram base para esta pesquisa como elementos condutores das análises, isto é, a informação neles registrada foi o norte para o estudo.

Contudo, houve casos também em que, mesmo nessas obras, a etimologia parecia obscura, ou controversa, e então foram consultadas outras fontes como referência, como foi com a confirmação da base etimológica de *caçuá*, que vigorava inicialmente como de origem tupi ou africana, nas obras

pesquisadas, e apenas uma consulta ao Grupo de Estudos de Línguas Africanas⁶² pode confirmar a origem banta (língua africana) da unidade léxica em pauta, excluindo-a desta pesquisa.

Assim, pode-se concluir, enfim, após as análises e considerações realizadas por este trabalho durante o período de sua execução que a variante brasileira do português abriga ainda a influência indígena em seu repertório vocabular, aqui representada pelas unidades léxicas de base indígena que foram auferidas como respostas para as 62 questões do Questionário Semântico-Lexical do projeto ALiB relacionadas às áreas semânticas das *atividades agropastoris*, da *fauna* e da *alimentação e cozinha* aqui analisadas.

E também se pode considerar que, mesmo essa presença sendo bem menor que em outros tempos, como na época da colonização do Brasil, em que a língua mais falada no país era de base indígena, os vestígios indígenas foram identificados e demonstraram profunda importância no recorte de dados apurados, pelos motivos já explicitados ao longo deste trabalho.

Pode-se concluir, finalmente, que a variante brasileira do português tem o seu desenhar tal qual outras línguas vivas, que se revelam a partir da renovação lexical promovida, dentre outros fatores motivadores, pelo intercâmbio sociocultural, neste caso demonstrado pelas unidades lexicais de diversas bases linguísticas aqui representadas.

O homem, sujeito agente desse intercâmbio cultural, sempre leva consigo por seus caminhos muito mais que pertences materiais, mas também sua história de vida e de seu grupo, sua cultura e suas tradições, que desenham sua língua. E, nesse sentido, “o léxico é o nível da língua que melhor evidencia as

⁶² O Grupo em questão foi contatado por meio de seu endereço eletrônico: gela@usp.br e é coordenado pela professora Margarida Maria Taddoni Petter.

pegadas do homem na sua trajetória histórica. [...] As migrações do homem se traduzem também em migrações de palavras” (ISQUERDO, 2009, p. 43).

REFERÊNCIAS

AGUILERA, Vanderci de Andrade. Os tupinismos na língua rural paranaense. *Línguas e Letras*. Vol. 7, n. 12, 1º semestre de 2006, p.99-125. Estudos Linguísticos. Disponível em: <http://e-revista.unioeste.br/index.php/linguaseletras/article/download/893/758>. Acesso em: 22 maio 2012.

ATLAS NACIONAIS. PROJETO ALIB. Disponível em: <http://twiki.ufba.br/twiki/bin/view/Alib/AtlasNacionais>. Acesso em: 05 nov 2012.

ANTONIL, André João. *Cultura e opulência do Brasil*. 3 ed. Belo Horizonte: Ed. Itatiaia; São Paulo: Ed. da Universidade de São Paulo, 1997.

BETHELL, Leslie (Org.). *História da América Latina: América Latina Colonial*. Volume I. Tradução Maria Clara Cescato. 2 ed. São Paulo: Editora da Universidade de São Paulo; Brasília: Fundação Alexandre de Gusmão, 2004.

BIDERMAN, Maria Teresa de Camargo. *Teoria lingüística: lingüística quantitativa e computacional*. Rio de Janeiro: Livros Técnicos e Científicos, 1978.

BIDERMAN, Maria Teresa de Camargo. O conhecimento, a terminologia e o dicionário. In: *Ciência e Cultura*, 58 (2), 2006. p. 35-37.

BIDERMAN, Maria Teresa de Camargo. Dimensões da palavra. In: *Filologia e Linguística Portuguesa*, n.2, p. 81-118, 1998. Disponível em: <http://www.fflch.usp.br/dlcv/lport/flp/images/arquivos/FLP2/Biderman1998.pdf>. Acesso em: 17 jun 2012.

BORGES, Luiz C. A instituição de línguas gerais no Brasil. In: ORLANDI, Eni P. (Org). *História das ideias linguísticas*. Cáceres: UNEMAT Editora, 2001. P. 199-222.

BRASIL ESCOLA. *Tratado de Tordesilhas*. Disponível em: <http://www.brasilecola.com/historiaab/tratado-de-tordesilhas.htm>. Acesso em: 12 jan 2012.

CAMACHO, Roberto Gomes. Sociolingüística II. In: BENTES, Anna Christina; MUSSALIM, Fernanda (Orgs.). *Introdução à lingüística: domínios e fronteiras*, v.1 – 5. ed. São Paulo: Cortez, 2005. P. 49-75.

CAMINHA, Pero Vaz de. *Carta a El Rei D. Manuel*. São Paulo: Dominus, 1963. Disponível em: <http://www.culturabrasil.org/zip/carta.pdf>. Acesso em: 15 dez 2011.

CARDOSO, Armando Levy. *Amerigenismos*. Tomo I. Rio de Janeiro: Biblioteca do Exército, 1961.

CARDOSO, Suzana Alice Marcelino. O Projeto ALiB e sua trajetória. In: MOTA, Jacyra Andrade; CARDOSO, Suzana Alice Marcelino (Orgs.). *Documentos 2: Projeto Atlas Linguístico do Brasil*. Salvador: Quarteto, 2006. P. 27-34.

CARDOSO, Suzana Alice Marcelino. Reflexões sobre a Dialectologia. In: ISQUERDO, Aparecida Negri (Org.). *Estudos geolingüísticos e dialetais sobre o português*. Campo Grande: Editora UFMS, 2008. p. 13-31.

CARDOSO, Suzana Alice Marcelino. *Geolingüística: tradição e modernidade*. São Paulo: Parábola Editorial, 2010.

COMITÊ NACIONAL DO PROJETO ALIB. *Atlas lingüístico do Brasil: questionário 2001*. Londrina: Eduel, 2001.

CORREIA, Margarita. Para a compreensão do conceito de 'empréstimo interno': primeira abordagem. In: ISQUERDO, Aparecida Negri; FINATTO, Maria José Bocorny. *As ciências do léxico: Lexicologia, Lexicografia, Terminologia*. Campo Grande: Ed. UFMS, 2008. p. 39-63.

COSERIU, Eugênio. *Fundamentos e tarefas da Sócio e da Etnolingüística*. I CONSEL. João Pessoa: 1978. (Mimeo).

COSERIU, Eugênio. *Lições de linguística geral*. Tradução: Evanildo Bechara. Rio de Janeiro: AO Livro Técnico, 1980.

COSERIU, Eugênio. *O homem e sua linguagem*. Trad. Carlos Alberto da Fonseca e Mário Ferreira. Coleção Linguagem. Rio de Janeiro/São Paulo: Presença/Edusp, 1982.

COSTA, Luiz Edmundo. *História da cachaça*. Disponível em: <http://www.brasilcult.pro.br/cachaca/historia.htm>. Acesso em: 18 set 2012.

CUNHA, Antônio Geraldo da. *Dicionário histórico das palavras portuguesas de origem tupi*. 3 ed. São Paulo: Companhia Melhoramentos, 1989.

CUNHA, Antônio Geraldo da. *Dicionário histórico das palavras portuguesas de origem tupi*. 4ª ed. São Paulo: Companhia Melhoramentos, 1998.

CUNHA, Antônio Geraldo da. *Dicionário etimológico da língua portuguesa*. 4 ed. Rio de Janeiro: Lexikon, 2010.

DANTAS, Beatriz G.; SAMPAIO, José Augusto L.; CARVALHO, Maria Rosária G. de. Os povos indígenas no Nordeste brasileiro. Um esboço histórico. In: CUNHA, Manuela Carneiro da (Org.). *História dos índios do Brasil*. São Paulo: Companhia das Letras; Secretaria Municipal de Cultura; FAPESP, 1992. P. 431-456.

DICIONÁRIO ONLINE DE LÍNGUA PORTUGUESA. Disponível em: <http://www.dicio.com.br/uca/>. Acesso em: 18 set 2012.

DICK, Maria Vicentina de Paula do Amaral. Toponímia e línguas indígenas no Brasil. In: *Estudos Avançados*, v.8, nº 22. São Paulo: USP, set/dez 1994, p.435-436.

DICK, Maria Vicentina de Paula do Amaral. Inter-relação léxico e cultura na América Indígena. Estudo de Caso. In: *Acta Semiótica et Lingüística (SBPL)*, São Paulo, v.8, p 295-308, 2000.

DICK, Maria Vicentina de Paula do Amaral. A variação do léxico em situação interétnica no português do Brasil. O indígena e o africano. In: *Anais do I Encontro de Filologia Românica*. São Paulo: FFLCH/USP, 2006.

DUBOIS, Jean *et al.* *Dicionário de lingüística*. [direção e coordenação geral da tradução Izidoro Blikstein]. São Paulo: Cultrix, 2006.

ERIKSON, Philippe. Uma singular pluralidade: a etno-história Pano. Tradução: Beatriz Perrone-Moisés. In: CUNHA, Manuela Carneiro da (Org.). *História dos índios do Brasil*. São Paulo: Companhia das Letras; Secretaria Municipal de Cultura; FAPESP, 1992. P. 239-252.

FAUSTO, Boris. *História do Brasil*. 13 ed. São Paulo: Editora da Universidade de São Paulo, 2008.

FAUSTO, Carlos. Fragmentos de história e cultura tupinambá. Da etnologia como instrumento crítico de conhecimento etno-histórico. In: CUNHA, Manuela Carneiro da (Org.). *História dos índios do Brasil*. São Paulo: Companhia das Letras; Secretaria Municipal de Cultura; FAPESP, 1992. P. 381-396.

FERNANDES, Florestan. Antecedentes indígenas: organização social das tribos tupis. In: HOLANDA, Sérgio Buarque de; CAMPOS, Pedro Moacyr. *Do descobrimento à expansão territorial*. 1º vol. Tomo 1: A época colonial. Coleção História geral da civilização brasileira. São Paulo: DIFEL, 1985. P. 72-86.

FERREIRA, Aurélio Buarque de Holanda. *Novo Dicionário Aurélio da Língua Portuguesa*, Versão 5.0, Curitiba: Editora Positivo, 2004.

FERREIRA, Carlota; CARDOSO, Suzana Alice Marcelino. *A dialetologia no Brasil*. São Paulo: Contexto, 1994.

GREGÓRIO, Ir. José. *Contribuição indígena ao Brasil. Lendas e tradições – usos e costumes – Fauna e flora – Língua – Raízes – Toponímia – VOCABULÁRIO*. III Volume. Belo Horizonte, MG: União Brasileira de Educação e Ensino, 1980.

GUIRAUD, Pierre. *La Semántique*. Paris: PUF, 1995.

HJELMSLEV, Louis. Langue et parole. 1943. In: *Essais Linguistiques*. Copenhague: Travaux du Cercle Linguistique de Copenhague, 1969. P. 29-44.

HOLANDA, Sérgio Buarque de. *Raízes do Brasil*. 20 ed. Rio de Janeiro: José Olympio, 1988.

HOUAISS, Antônio. *O português no Brasil*. Rio de Janeiro: UNIBRADE – Centro de Cultura, 1985.

HOUAISS, Antônio. *Dicionário Eletrônico Houaiss da Língua Portuguesa*, Versão 1.0. Rio de Janeiro: Editora Objetiva, 2001.

IBGE. Censo Demográfico 2000 – resultados da amostra. Disponível em: http://www.ibge.gov.br/home/estatistica/populacao/censo2000/primeiros_resultados_amostra/grandes_regioes/pdf/tabela_2_1_1.pdf. Acesso em: 23 maio 2012.

INFOESCOLA. *Tratado de Madri de 1750*. Disponível em: <http://www.infoescola.com/historia/tratado-de-madrid-de-1750/>. Acesso em: 12 jan 2012.

ISQUERDO, Aparecida Negri. Léxico em tempo e espaço: a questão dos regionalismos. In: MARIN, Jérri Roberto; VASCONCELOS, Cláudio Alves de (Orgs.). *História, região e identidades*. Campo Grande, MS: Ed. UFMS, 2003. P. 165-181.

ISQUERDO, Aparecida Negri. Atlas regionais em andamento no Brasil: perspectivas metodológicas. In: AGUILERA, Vanderci de Andrade (Org.). *A geolinguística no Brasil – trilhas seguidas, caminhos a percorrer*. Londrina: Eduel, 2005. P. 334-356.

ISQUERDO, Aparecida Negri. Achega para a discussão do conceito de regionalismos no Português do Brasil. In: *Alfa: Revista de Lingüística / UNESP* v. 50, nº 2. São Paulo: UNESP, 2006. P. 9-24.

ISQUERDO, Aparecida Negri. O caminho do rio, o caminho do homem, o caminho das palavras... In: RIBEIRO, Silvana Soares Costa; COSTA, Sônia Bastos Borba; CARDOSO, Suzana Alice Marcelino. *Dos sons às palavras*. Nas trilhas da língua portuguesa. Salvador: EDUFBA, 2009, p. 42-59.

LARA, Luis Fernando. *El concepto de norma em Lingüística*. México: El Colegio de México, 1976.

LINGUAMÓN. CASA DAS LÍNGUAS. *Mapudungu*. Disponível em: http://www10.gencat.cat/pres_casa_llengues/AppJava/frontend/llengues_detall.jsp?id=260&idioma=16. Acesso em: 03 out 2012.

LUCCHESI, Dante. *Sistema, mudança e linguagem: um percurso da linguística neste século*. Lisboa: Colibri Artes gráficas, 1998.

LUCCHESI, Dante; BAXTER, Alan; RIBEIRO, Ilza (Orgs.). *O português afro-brasileiro*. Salvador: EDUFBA, 2009.

MACHADO, José Pedro. *Dicionário etimológico da Língua Portuguesa*. Vol. 2 e 4. Lisboa: Livros Horizontes Ltda., 1987.

MARIANI, Bethânia. A institucionalização da língua, história e cidadania no Brasil do século XVIII: o papel das academias literárias e da política do Marquês de Pombal. In: ORLANDI, Eni P. (Org.). *História das ideias linguísticas*. Cáceres: UNEMAT Editora, 2001. P. 99-124.

MATTOSO CÂMARA JUNIOR, Joaquim. *Introdução às línguas indígenas brasileiras*. 3 ed. Rio de Janeiro: AO Livro Técnico S/A, 1977.

MATTOSO CÂMARA JUNIOR, Joaquim. *Dicionário de Linguística e gramática referente à língua portuguesa*. Petrópolis: Vozes, 1998.

MATTOS E SILVA, Rosa Virgínia. *Ensaio para uma sócio-história do português brasileiro*. 2 ed. São Paulo: Parábola Editorial, 2004.

MEADER, Robert E. Índios do Nordeste. Levantamento sobre os remanescentes tribais do nordeste brasileiro. Tradução: Yonne Leite. Disponível em: <http://www.sil.org/americas/brasil/publcns/ling/indnord.pdf>. Acesso em 19 jun 2012.

MELATTI, Júlio Cezar. *Índios do Brasil*. 3 ed. São Paulo: HUCITEC; Brasília: INL, 1980.

MELO, Gladstone Chaves de. *A língua do Brasil* – 4 ed. Melhorada e aumentada. Rio de Janeiro: Padrão, 1981.

MENÉNDEZ, Miguel A. A área Madeira-Tapajós. Situação de contato e relações entre colonizador e indígena. In: CUNHA, Manuela Carneiro da (Org.). *História dos Índios do Brasil*. São Paulo: Companhia das Letras; Secretaria Municipal de Cultura; FAPESP, 1992. P. 281-296.

MOTA, Jacyra Andrade; CARDOSO, Suzana Alice Marcelino. Para uma nova divisão dos estudos dialetais brasileiros. In: _____ (Orgs.). *Documentos 2: projeto Atlas Linguístico do Brasil*. Salvador: Quarteto, 2006. P. 15-26.

MOUTON, Pilar García. Dialectología y Geografía Lingüística. In: ALVAR, Manuel (Org.). *Manual de Dialectología hispánica*. El Español de España. Barcelona: Editora Ariel, 2009. P. 63-77.

MUSEU NACIONAL. Disponível em: <http://www.museunacional.ufrj.br/linguistica/historico.htm>. Acesso em 19 jun 2012.

NASCENTES, Antenor. *Dicionário etimológico resumido*. Rio de Janeiro: Instituto Nacional do Livro, Ministério Educação e Cultura, 1966. Coleção dicionários especializados.

NAVARRO, Eduardo de Almeida. *Método moderno do tupi antigo: a língua do Brasil nos primeiros séculos*. Petrópolis: Vozes, 1998.

NELI, Hervé Théry; MELLO, Aparecida de. *Atlas do Brasil – disparidades e dinâmicas do território*. São Paulo: Editora da USP, 2005.

OLIVEIRA, Ana Maria Pinto Pires de. *O português do Brasil: brasileiros e regionalismos*. Tese de doutoramento. Araraquara: UNESP, Faculdade de Ciências e Letras, 1999.

PETTER, Margarida. Africanismos no português do Brasil. In: ORLANDI, Eni P. (Org.). *História das ideias linguísticas*. Cáceres: UNEMAT Editora, 2001. P. 223-234.

PRADO JÚNIOR, Caio. *Formação do Brasil contemporâneo*. 19 ed. São Paulo: Brasiliense, 1986.

PROJETO ATLAS LINGUÍSTICO DO BRASIL. *Destaques*. Disponível em: <http://twiki.ufba.br/twiki/bin/view/Alib/WebHome>. Acesso em: 17 abr 2012.

RAMIREZ, Arnulfo G. Dialectologia y Sociolingüística. In: ALVAR, Manuel (Org.). *Manual de Dialectologia hispânica*. El Español de España. Barcelona: Editora Ariel, 2009. P. 37-48.

REIS, Regiane Coelho Pereira. *Atlas lingüístico do município de Ponta Porã-MS: um registro das línguas em contatos na fronteira do Brasil com o Paraguai*. Dissertação de Mestrado. Universidade Federal de Mato Grosso do Sul, 2006.

RIBEIRO, Darcy. *O povo brasileiro*. A formação e o sentido do Brasil. São Paulo: Companhia das Letras, 1995.

RODRIGUES, Aryon Dall'igna. *Línguas brasileiras: para o conhecimento das línguas indígenas*. 4 ed. São Paulo: Edições Loyola, 2002.

RODRIGUES, Aryon Dall'igna. *Sobre as línguas indígenas e sua pesquisa no Brasil*. São Paulo: Ciência e cultura, vol. 57 n. 02. Abr/Jun 2005. P. 35-38.

RODRIGUES, Aryon Dall'igna. *A originalidade das línguas indígenas brasileiras*. Disponível em: <http://www.comciencia.br/reportagens/linguagem/ling13.htm>. Acesso em: 17 abr 2012a.

RODRIGUES, Aryon Dall'igna. As línguas gerais sul-americanas: In: *Papia*: Revista Brasileira de Estudos Crioulos e Similares, vol. 4, n.2, 1996. Disponível em: <http://www.abecs.net/ojs/index.php/papia/article/view/143/210>. Acesso em: 04 set 2012b.

ROSSI, Nelson. *Atlas Prévio dos Falares Baianos*. Rio de Janeiro: Instituto Nacional do Livro/Ministério da Educação e Cultura, 1963.

SALVADOR, Frei Vicente de. *História do Brasil. Em que se trata do Descobrimento do Brasil, costumes dos naturais, aves, peixes, animais e do mesmo Brasil*. Salvador, 1627. Disponível em: <http://www.dominiopublico.gov.br/download/texto/bn000138.pdf>. Acesso em: 15 dez 2011.

SAMPAIO, Teodoro. *O tupi na geographia nacional*. Salvador: Secção Graphica da Escola de Aprendizes Artificies, 1928.

SAPIR, Edward. *Linguística como ciência*. Rio de Janeiro: Editora Livraria Acadêmica, 1969.

SAPIR, Edward. *A Linguagem: Introdução ao estudo da fala*. Tradução Joaquim Mattoso Câmara Júnior. 2 ed. Rio de Janeiro: Acadêmica, 1971. P. 205-216.

SAUSSURE, Ferdinand de. *Curso de Linguística Geral*. São Paulo: Cultrix, 2006.

SEKI, Lucy. *Línguas indígenas do Brasil no limiar do século XXI*. Disponível em: <http://www.unimep.br/phpg/editora/revistaspdf/imp27art11.pdf>. Acesso em: 15 ago 2011.

SILVA, Antonio Moraes. *Diccionario da lingua portugueza* - recompilado dos vocabulários impressos até agora, e nesta segunda edição novamente emendado e muito acrescentado, por ANTONIO DE MORAES SILVA. Lisboa: Typographia Lacerdina, 1813.

SILVA, Greize Alves; AGUILERA, Vanderci de Andrade. Geolinguística: um estudo no campo lexical da fauna brasileira. In: *Anais do V Encontro Científico do Curso de Letras. O desafio das Letras*. 2007. Disponível em: http://www.faccar.com.br/desletras/hist/2007_g/textos/14.htm. Acesso em: 22 maio 2012.

SILVA, Lidiane Martins da; AGUILERA, Vanderci de Andrade. Variantes lexicais para o *gambá*: um estudo geolinguístico nas capitais brasileiras. In: *1º CIELLI – Colóquio Internacional de Estudos Linguísticos e Literários*, 2010, Maringá. Anais. Maringá: Universidade Estadual de Maringá, 2010, p. 1-13. Disponível em: <http://www.cielli.com.br/downloads/529.pdf>. Acesso em: 03 out 2012.

TAMBA, Irène. *A Semântica*. São Paulo: Parábola Editorial, 2006.

TAVARES, Marineide Cassuci. *Estudo toponímico da região centro-norte de Mato Grosso do Sul: o desvendar de uma história*. Dissertação (Mestrado). Três Lagoas: Universidade Federal de Mato Grosso do Sul, 2005.

TIBIRIÇÁ, Luiz Caldas. *Dicionário de topônimos brasileiros de origem tupi. Significação dos nomes geográficos de origem tupi*. São Paulo: Traço Editora, 1997.

URBAN, Gred. A história da cultura brasileira segundo as línguas nativas. Tradução: Beatriz Perrone-Moisés. In: CUNHA, Manuela Carneiro da (Org.). *História dos Índios do Brasil*. São Paulo: Companhia das Letras; Secretaria Municipal de Cultura; FAPESP, 1992. P. 87-102.

VIANNA, Hélio. *História do Brasil*. 3 ed. São Paulo: Edições melhoramentos, 1965a. Volume I – Período colonial.

VIANNA, Hélio. *História do Brasil*. 3 ed. São Paulo: Edições melhoramentos, 1965b. Volume II – Monarquia e República.

WIKIPÉDIA. Disponível em: http://pt.wikipedia.org/wiki/P%C3%A3o_franc%C3%AAs. Acesso em: 15 ago 2012.

WRIGHT, Robin W. História indígena do noroeste da Amazônia. Hipótese, questões e perspectivas. In: CUNHA, Manuela Carneiro da (Org.). *História dos índios do Brasil*. São Paulo: Companhia das Letras; Secretaria Municipal de Cultura; FAPESP, 1992. P. 253-266.