REVISTA DA
UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA

SITIENTIBUS

ESTUDOS LINGÜÍSTICOS E FILOLÓGICOS

ARTIGO

ESTUDOS SOBRE O LÉXICO DE ORIGEM INDÍGENA NO SEMIÁRIDO BAIANO

STUDIES ON THE LEXICON OF INDIGENOUS ORIGIN IN THE SEMI-ARID REGION OF BAHIA

SAÁDIA RAMOS FERREIRA

Mestranda em Estudos Linguísticos (PPGEL/UEFS). E-mail: ferreirasaadia@gmail.com

NORMA LUCIA FERNANDES DE ALMEIDA

Doutora em Linguística pela Universidade Estadual de Campinas (Unicamp). Professora do Departamento de Letras e Artes da UEFS
E-mail: norma@uefs.br

RESUMO

O léxico pode ser entendido como um conhecimento compartilhado por falantes que se formam no âmbito do saber vocabular de um grupo sociolinguístico cultural. Além disso, o léxico mantém uma estreita relação com a história cultural de uma comunidade. O estudo do léxico é feito com base em três principais disciplinas: a lexicologia, a lexicografia e a terminologia. A pesquisa apresentada neste artigo teve como objetivo analisar as contribuições das línguas indígenas na formação do Português Brasileiro (PB) e, para isso, foram utilizados os quatro volumes que compõem o *corpus* Amostras da Língua Falada no Semiárido Baiano (Almeida; Carneiro, 2008), a fim de buscar lexias que possuam origem nas línguas indígenas, do tronco Tupi e do tronco Macro-Jê. Os dados identificados no material analisado mostram o quanto as línguas indígenas contribuíram e deixaram marcas até os dias de hoje no léxico utilizado no semiárido baiano e, consequentemente, no Português Brasileiro.

Palavras-chave: Léxico. Lexicografia. Semiárido. Tupi. Macro-Jê.

ABSTRACT

The lexicon can be understood as knowledge shared by speakers that is formed within the scope of the vocabulary knowledge of a sociolinguistic cultural group. Furthermore, the lexicon maintains a close relationship with the cultural history of a community. The study of the lexicon is based on three main disciplines: lexicology, lexicography and terminology. The research presented in this article aimed to analyze the contributions of indigenous languages in the formation of Brazilian Portuguese (BP). To this end, the four volumes that make up the corpus *Amostras da Língua Falada no Semiárido Baiano* (Almeida; Carneiro, 2008) were used in order to search for lexicons that have their origin in indigenous languages, from the Tupi and Macro-Jê trunk. The data identified in the analyzed material show how much indigenous languages contributed and left their mark to this day on the lexicon used in the semi-arid region of Bahia and, consequently, on Brazilian Portuguese.

Keywords: Lexicon. Lexicography. Semi-arid. Tupi. Macro-Jê.

1 INTRODUÇÃO

O léxico pode ser conceituado como um conjunto infinito de palavras de um determinado idioma, que carrega consigo as representações que um grupo, de determinada língua, faz do mundo. O léxico do português brasileiro foi formado, principalmente, pela incorporação de lexias indígenas e africanas, a partir da vinda dos colonizadores para essas terras, trazendo com eles africanos escravizados.

Apesar do Tupi Antigo não ser mais uma língua plenamente utilizada como era até o período colonial, deixou fortes traços no léxico do português brasileiro. Nas primeiras décadas de contato europeu com língua da costa, os colonizadores portugueses enviavam correspondentes para o Brasil, para que eles aprendessem a língua mais falada na costa do país e quando as expedições chegassem, tais pessoas servissem como guias (Rodrigues, 1993).

O semiárido da Bahia também possuiu, e ainda possui, comunidades indígenas, portanto, é com base nesse pressuposto que surgiu o interesse para a realização desta pesquisa. O intuito foi investigar o que ainda é utilizado na variedade do português falada no semiárido baiano que possui origem nos troncos linguísticos Tupi e Macro-Jê, uma vez que “há, como se sabe, vários empréstimos lexicais das línguas de base tupi incorporados no português brasileiro, embora muitas vezes não se tenha consciência disso” (Ivo, 2021, p.49).

Este trabalho foi realizado tendo como base os quatro volumes que compõem o *corpus Amostras da Língua Falada no Semiárido Baiano*, com a finalidade de investigar a ocorrência das lexias de base indígena, de diferentes troncos linguísticos, em diversas comunidades do semiárido da Bahia. Fizemos o levantamento das ocorrências dos vocábulos de possível origem indígena e, a partir desse levantamento, foram pesquisadas suas respectivas ocorrências em dicionários antigos e contemporâneos da língua portuguesa.

Os estudos acerca das línguas indígenas não são recentes, porém, é importante que possamos conhecer mais a respeito, sendo de extrema relevância para o resgate, para o conhecimento das línguas dos povos originários e para a conscientização de quem foram os primeiros habitantes das terras hoje chamadas de Brasil (Cunha e Souza, 2011), demonstrando que, apesar de tantas transformações e tentativas de apagamento ao longo do tempo, ainda são encontrados grandes vestígios e contribuições das línguas dos povos originários no léxico atual da língua portuguesa falada no Brasil.

2 DESCOBRINDO O LÉXICO

O léxico pode ser definido como o conjunto infinito de palavras utilizadas por falantes de uma determinada língua que carrega consigo as representações que um grupo, de determinada língua ou dialeto, faz do mundo. Oliveira e

Isquerdo (2001, p. 9) reiteram que o léxico é o primeiro meio de acesso a um texto, podendo ser considerado como uma janela pela qual “uma comunidade pode ver o mundo”. Além disso, Oliveira e Isquerdo (2001) apresentam o léxico como um conhecimento compartilhado por falantes que se formam no âmbito do saber vocabular de um grupo sociolinguístico cultural, além de manter uma estreita relação com a história cultural da comunidade e relacionar-se com a nomeação e com a compreensão, com isso, um indivíduo estará sempre vinculado a representar apenas as estruturas linguísticas da sua língua materna e não as estruturas de uma L2, por exemplo.

Barreiros (2017, p. 112) afirma que “o léxico constitui um amplo universo conceitual, devido à capacidade do falante de usar, de criar e de renovar a sua língua”. Por meio do léxico é possível perceber as crenças, os valores, os hábitos e costumes de uma dada comunidade, sobre isso Oliveira e Isquerdo (2001, p. 9) afirmam que “na medida em que o léxico recorta realidades do mundo, define, também, fatos de cultura”. Neste trabalho, levamos em consideração o léxico numa perspectiva sociocultural já que um dos nossos objetivos é contribuir com projetos de revitalização de línguas indígenas e, conseqüentemente, da cultura dos povos indígenas no semiárido baiano.

Para Biderman (1981, p. 138) o léxico é “como o tesouro vocabular de uma determinada língua. Ele inclui a nomenclatura de todos os conceitos linguísticos e não linguísticos e de todos os referentes do mundo físico e do universo cultural [...]”. O patrimônio lexical de uma comunidade evidencia suas singularidades culturais, visto que esse nível da língua reflete as crenças e valores de um dado grupo. O léxico também pode ser considerado como um vasto campo de conhecimento linguístico. Antunes (2012, p. 29) fala que

O léxico, ao contrário, é aberto, inesgotável, constantemente renovável, não apenas porque surgem novas palavras, mas, também pela dinâmica interna das palavras, que vão e vêm, que desaparecem e reaparecem, que mantêm seus significados ou os mudam, de um lugar para outro, de um tempo para outro.

A partir desse movimento que acontece com a língua, palavras podem passar a ser utilizadas no léxico, enquanto outras que costumavam ser usadas, deixam de ser. Antunes (2012) caracteriza esses movimentos como lexicalização e deslexicalização. Para a autora, a lexicalização é quando ocorre o processo do surgimento de uma nova palavra no léxico utilizado por uma comunidade. Já a deslexicalização é quando ocorre o contrário, uma palavra que costumava ser utilizada no vocabulário de uma comunidade cai em desuso ou passa a ser utilizada em outras situações mais limitadas. São processos que nós acreditamos que ocorreram com o PB a partir do contato com as línguas indígenas e africanas.

O léxico, pois, é diferente de vocabulário. De acordo com Vilela (1995, p. 13) léxico é o conjunto de palavras que membros de uma comunidade comunicam entre si, já o vocabulário é o conjunto de vocábulos que realmente existem em um determinado lugar ou tempo.

Os estudos sobre o léxico de uma língua são desenvolvidos a partir de três principais disciplinas que são a lexicologia, a lexicografia e terminologia. A lexicologia é uma ciência antiga e que se responsabiliza pelo estudo e análise das palavras. Ela também pode ser conceituada como o estudo científico do léxico (Vilela, 1995, p. 14). Para Abreu (2019, p. 38) "o léxico de uma língua representa a comunidade que o utiliza, com isso podemos dizer que a lexicologia é a área do conhecimento que estuda o léxico a partir de seus significados e de suas representações", ou seja, o estudo realizado pelos lexicólogos parte de entender, inicialmente, o significado até chegar ao significante. E para que se possa compreender da melhor forma os significados das lexias, é necessário conhecer sobre a cultura e as representações que a comunidade faz a respeito do léxico a qual utiliza.

A arte ou técnica da elaboração de dicionários, glossários e vocabulários cabe à ciência lexicográfica. Essa disciplina do léxico é antiga, em se tratando da elaboração de dicionários. De acordo com Biderman (2001, p. 17), a lexicografia começou nos séculos XVI e XVII, a partir da elaboração de dicionários em língua portuguesa, o primeiro chama-se *Vocabulário Português-Latino* de Rafael Bluteau e o segundo *Dicionário da Língua Portuguesa* de Antônio de Moraes Silva. A língua apresenta uma linguagem geral, que é o que estuda a lexicografia, e em determinadas situações apresenta uma linguagem mais especializada, como linguagens técnicas, profissionais, regionais, científicas e de grupos sociais diferentes. Cabe à terminologia o estudo dos termos específicos de cada situação citada acima. Essa pode, facilmente, ser confundida com o conceito da lexicologia, porém, a lexicologia se encarrega de estudar todos os termos pertencentes ao léxico de uma língua, já a terminologia estuda as palavras que constituem linguagens mais especializadas, ou seja, a linguagem própria de um determinado grupo.

As linguagens especializadas se caracterizam pelo emprego da terminologia, que representa a estrutura conceptual de determinada matéria, enquanto os termos denominam os conceitos da rede estruturada da matéria em questão. A diferença fundamental entre um texto da língua geral e outro, de uma linguagem especializada, está no uso dos termos específicos de determinada área, que lhe confere o caráter de especificidade, em distintos níveis de especialização, conforme o tipo de matéria e seu grau de abstração (Andrade, 2001, p. 193).

Com isso, podemos dizer que a lexicografia também é responsável por representar a cultura e a identidade de um povo através dos registros dos significados dos verbetes nas obras lexicográficas. "A lexicografia não se reduz a um

problema semântico, embora o acolha, uma vez que o sentido que interessa à lexicografia é definido nos usos socioculturais das palavras" (Santos, 2016, p. 143).

3 PERCURSO METODOLÓGICO

O *corpus* utilizado para obtenção de dados para a realização desta pesquisa foi a "Coleção Amostras da Língua Falada no Semi-árido Baiano" – que é formado por quatro volumes, somando 72 entrevistas gravadas em algumas comunidades do semiárido da Bahia e foi constituído entre os anos de 1994 a 2002 pelas professoras Norma Lúcia Fernandes de Almeida e Zenaide de Oliveira Novais Carneiro –, quando se investigou a ocorrência de lexias de base indígena, de diferentes troncos linguísticos, em algumas comunidades do semiárido baiano. Para a análise do *corpus*, seguimos os princípios da lexicografia moderna com o auxílio do programa *AntConc* para selecionar as lexias a serem investigadas. Foi levada em consideração a proposta de Sardinha (2000) sobre a linguística de *corpus*, que foi utilizada para identificação e análise das lexias selecionadas através de um programa computacional.

A Linguística de Corpus ocupa-se da coleta e exploração de corpora, ou conjuntos de dados linguísticos textuais que foram coletados criteriosamente com o propósito de servirem para a pesquisa de uma língua ou variedade linguística. Como tal, dedica-se à exploração da linguagem através de evidências empíricas, extraídas por meio de computador (Sardinha, 2000, p. 325).

Fizemos um levantamento das ocorrências dos vocábulos de possível origem indígena nas entrevistas dos quatro volumes que compõem a coleção, com o auxílio do programa *Antconc*. Após isso, selecionamos uma lista de palavras para serem analisadas e, a partir desse levantamento, pesquisamos suas respectivas ocorrências em dicionários antigos (o *Novo Dicionário da Língua Portuguesa*, de Aurélio (1986), o *Aulete Digital*, o *Dicionário Tupi (antigo) – Português*, de Carvalho (1987), o *Dicionário Tupi Antigo*, de Navarro (2013) e o texto *Notas Sobre o Sistema de Parentesco dos Índios Kiriri*, de Rodrigues (1948) e contemporâneos da língua portuguesa, atestando-se que possuem origem indígena, que é o que buscamos, concluiu-se o processo de análise.

4 RESULTADOS OBTIDOS

Diante da pesquisa realizada, foi possível perceber que o tronco linguístico Tupi foi o que mais obteve ocorrências de lexias, comparado ao tronco Macro-jê. Esse fato pode ser explicado levando em consideração todo o contexto histórico da ocupação do território baiano pelos povos indígenas, onde houve uma grande quantidade de grupos falantes de famílias linguísticas do tronco Tupi, inclusive quando houve a instituição da Língua Geral Brasileira, e os poucos grupos de falantes de famílias linguísticas do tronco Macro-jê, acabaram

perdendo a língua materna com o passar do tempo e dos acontecimentos. Nesta pesquisa só foi encontrada a lexia *quixabeira* com possível origem no tronco Macro-jê, devido à região em que ela possui registros.

No volume I do *corpus*, constituído na zona rural de Anselino da Fonseca, foram registradas 19 lexias de origem indígena em uma pesquisa feita nos doze inquéritos que compõem o volume. Algumas dessas lexias foram aipim, baraúna, cajazeira, jacutinga, juriti, manaíba, paraná e teiú. A maioria das lexias apresentam, em seu contexto de uso, significações semelhantes às apresentadas pelos dicionários, com a exceção de algumas que diferem do significado original por, muitas vezes, denominarem localidades, transformando-se em um substantivo próprio. Ainda há aquelas lexias que no contexto de uso aparece com o significado apresentado pelos dicionários e com significações divergentes, como é o caso da lexia "Juazeiro".

No volume II do *corpus*, gravado na zona rural de Rio de Contas, foram registradas 16 lexias de base indígena nos vinte e quatro inquéritos gravados nas três comunidades, como: abaré, araponga, macaxeira, mangabeira, pequi, suçuarana e tamanduá. A única que aparece com o significado divergente é "ABARÉ", que os dicionários apresentam como sendo o nome dado por indígenas ao que conhecemos como

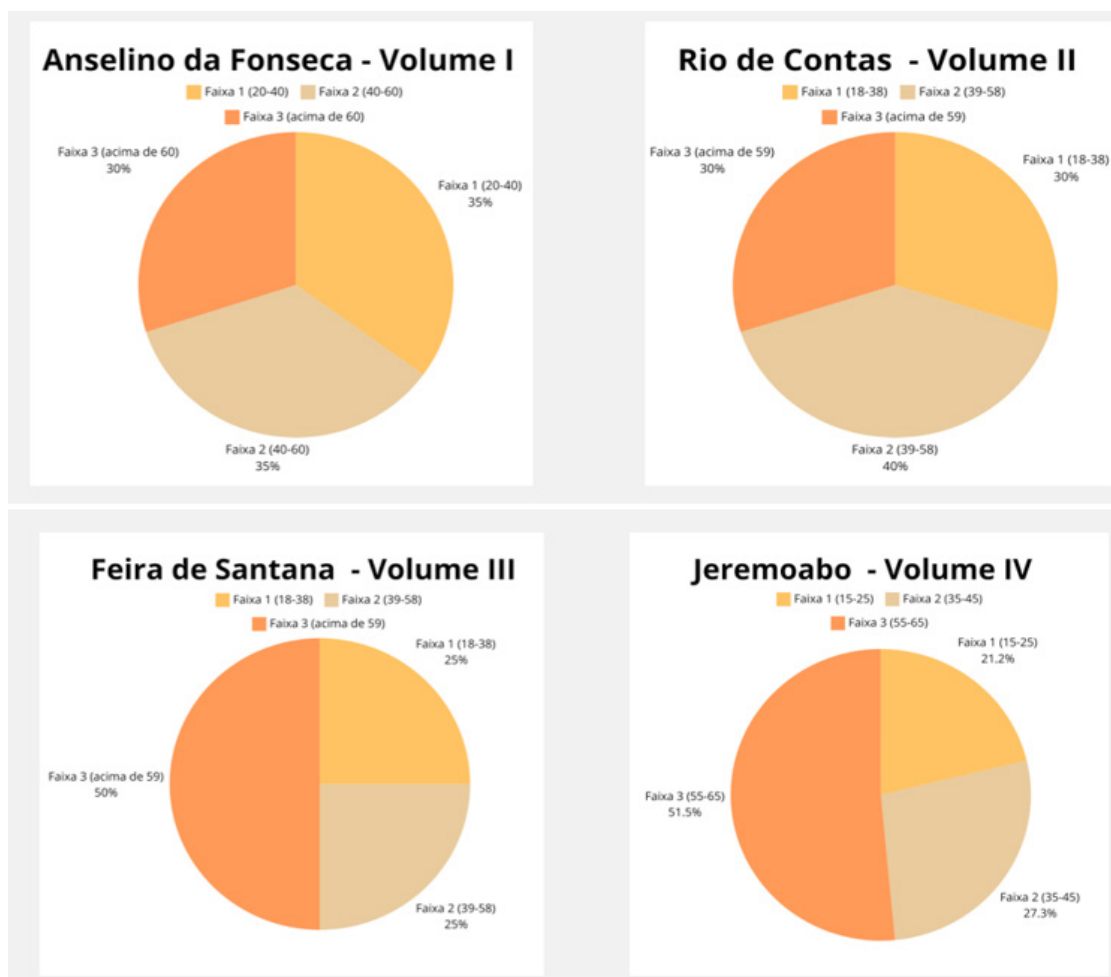
missionário, mas que nos inquéritos aparece nomeando uma cidade.

No volume III, que foi gravado na zona rural de Feira de Santana, foram registradas 8 lexias de base indígena nos inquéritos que compõem o *corpus*. Dentre elas estão: beiju, cajá, licuri, pitanga e tapioca. Todas as lexias analisadas possuem a mesma significação apresentada pelos dicionários em seus usos nos diálogos.

No volume IV, que foi gravado na zona rural de Jeremoabo, foram registradas 17 lexias de base indígena nos trinta e seis inquéritos analisados e que foram gravados em três comunidades da região, que são: arapuca, caititu, caju, jatobá, macambira, pindoba, tapera e umbuzeiro¹. Apenas uma lexia das que foram localizadas não está com a mesma significação apresentada pelos dicionários, que é "tapera". Nos inquéritos aparece nomeando uma localidade. Além de "quixabeira" que não foi possível localizar nos dicionários de Tupi, mas que há estudos que apontam que essa lexia está vinculada à família linguística Kiriri, que pertence ao tronco Macro-jê. Se assim for, temos apenas esta lexia no trabalho que possui origem no Macro-jê.

Os gráficos abaixo demonstram as faixas etárias dos informantes das entrevistas de cada volume do *corpus* e a porcentagem de lexias encontradas por faixa:

Figura 1 - Gráficos das faixas etárias dos entrevistados e a porcentagem de lexias encontradas



Conforme os resultados apresentados, foi possível perceber que, embora tantos anos já se passaram desde o período da colonização, ainda se encontra muitas palavras oriundas das línguas indígenas que já existiram e algumas que ainda existem (e resistem) na língua portuguesa falada atualmente. Algumas dessas línguas são pertencentes ao tronco Tupi e outras ao tronco Macro-jê. Com os dados obtidos após a realização deste estudo, encontramos mais lexias do tronco tupi, do que do tronco Macro-jê. Isso pode ser explicado pelo fato de que a língua adotada para comunicação pelos portugueses no período de colonização, ter sido do tronco Tupi e ter havido bastante expansão, também porque, segundo Pinto (1935, p. 147) os dados relacionados às línguas do tronco Macro-jê são bastante escassos e não há muitos registros das línguas pertencentes a esse tronco.

5 PALAVRAS FINAIS

Diante do que foi pesquisado, podemos perceber que as línguas indígenas, não mais faladas como língua materna, não foram totalmente apagadas, pois há muitas heranças delas na língua portuguesa que é utilizada nos tempos atuais no país, sendo um dos motivos que fazem com que a nossa língua portuguesa possua características diferentes da língua portuguesa deixada (e imposta) pelos colonizadores. Tânia Lobo, em seu artigo intitulado *Rosa Virgínia Mattos e Silva e a História Social Linguística do Brasil*, apresenta uma proposição de Mattos e Silva, onde a autora afirma que "a história linguística do Brasil não se restringe à história da língua portuguesa no Brasil, nem à história do português brasileiro" (2015, p. 71). A língua portuguesa existente hoje no Brasil carrega muitas marcas e heranças das línguas indígenas, como também de outras línguas que formaram o "multilinguismo" no território brasileiro no século XVIII, como as línguas africanas. Portanto, é possível – e cabível – afirmar que a história da língua que se fala no Brasil, atualmente, começa muito antes dessa língua existir nesse território. Conforme o que foi exposto neste trabalho, percebe-se que o que aconteceu para o estabelecimento da língua portuguesa no Brasil, foi, realmente, uma violência, tanto linguística, como física e cultural. Portanto, é importante e necessário o resgate, que vem sendo feito, das línguas indígenas, mesmo sabendo que é muito difícil o uso destas se abranger por todo o território, para que não se perca a história dos povos e a cultura dos povos originários, que integram também a história linguística e cultural do nosso país.

REFERÊNCIAS

ABREU, Uana Vanessa Pinheiro de. **Da África à Bahia: um estudo sobre o léxico africano em comunidades do semiárido baiano**. Dissertação. 2019.

ALMEIDA, Norma Lucia F. de; CARNEIRO, Zenaide de O. N. **Amstras da língua falada no semi-árido baiano**. Feira de Santana, UEFS/FAPESB, 2008.

ANDRADE, Maria Margarida de. Lexicologia, terminologia: definições, finalidades, conceitos operacionais. In: OLIVEIRA, Ana Maria Pires de; ISQUERDO, Aparecida Negri. **As ciências do léxico**: Lexicologia, Lexicografia, Terminologia. 2 ed. Campo Grande (MS): Ed. UFMS, 2001 [1998].

ANTUNES, Irlandé. O léxico de uma língua. In: _____. **O território das palavras**: estudo do léxico em sala de aula. São Paulo: Parábola Editorial, 2012.

BARBOSA, M. A. Lexicologia, lexicografia, terminologia, terminografia, identidade científica, objeto, métodos, campos de atuação. In **Anais do II Simpósio Latino-Americano de Terminologia**. I Encontro Brasileiro de Terminologia Técnico-Científica. Brasília, 1990.(p. 152-158).

BARREIROS, Liliene Lemos Santana. **Vocabulário de Eulálio Motta** / Liliene Lemos Santana Barreiros. Salvador, 2017.

BIDERMAN, M. T. C. **A ciência da Lexicografia**. In: ALFA, São Paulo, 28, 1984.

BIDERMAN, Maria Tereza C. A estrutura mental do léxico. In: BORBA, Francisco da Silva. (Org.). **Estudos de Filologia e Linguística**: em homenagem a Isaac Nicolau Salum. São Paulo: T.A Queiroz/Edusp, v. 02, 1981, p. 131-145.

BIDERMAN, Maria Tereza Camargo. As ciências do Léxico. In: ISQUERDO, Aparecida Negri; OLIVEIRA, Ana Maria Pinto Pires de. (Orgs). **As ciências do léxico**: lexicologia, lexicografia, terminologia. 2 ed. Campo Grande: Ed UFMG, 2001, p. 13-22. CARVALHO, Moacyr Ribeiro de. **Dicionário tupi (antigo) português**/ Moacyr Ribeiro de carvalho – Salvador, 1987.

CUNHA E SOUZA, Hirão Fernandes. **O português Kiriri**: aspectos fônicos e lexicais na fala de uma comunidade do sertão baiano / Hirão Fernandes Cunha e Souza. 2011.

FERREIRA, Aurelio Buarque de Holanda; **Novo Dicionário da Língua Portuguesa**. 2ª ed. J.E.M.M. Editores Ltda. - 1986, (p. 1-1821).

IRIARTE SANROMÁN, Á. **A Unidade Lexicográfica**. Palavras, Colocações, Frasesmas, Pragmatemas. Braga: Centro de Estudos Humanísticos-Universidade do Minho, 2001.

IVO, Ivana Pereira. O falar caiçara: subsídios para os estudos sobre a contribuição de línguas indígenas para a formação do português brasileiro. **Estudos da Língua(gem)**, [S. l.], v. 19,

¹Quanto à lexia “umbuzeiro”, já foi encontrada sendo utilizada por falantes de língua Macro-jê, porém tanto Navarro (2013) quanto Carvalho (1987) a apresentam como herança do Tupi Antigo.

n. 3, p. 29-53, 2021. DOI: 10.22481/el.v19i3.9139. Disponível em: <https://periodicos2.uesb.br/index.php/estudosdalinguagem/article/view/9139>. Acesso em: 29 out. 2024.

LOBO, T. **Rosa Virgínia Mattos e Silva e a história social linguística do Brasil**. Estudos de Linguística Galega, v. 7, 2015, p. 69-82.

NAVARRO, Eduardo de Almeida. **Dicionário de Tupi Antigo: a língua indígena clássica do Brasil**. São Paulo: Global, 2013.

OLIVEIRA, Ana Maria Pires de; ISQUERDO, Aparecida Negri. **As ciências do léxico: Lexicologia, Lexicografia, Terminologia**. 2 ed. Campo Grande (MS): Ed. UFMS, 2001[1998], (p. 9-123).

PINTO, Estêvão. **Os Indígenas do Nordeste**. Companhia Editora Nacional, São Paulo, 1935.

RODRIGUES, Aryon Dall'Igna. **Línguas Indígenas: 500 anos de descobertas e perdas**. D.E.L.T.A., vol. 9, n. 1, 1993, (p. 83-103).

RODRIGUES, Aryon Dall'Igna. **Notas sobre o sistema de parentesco dos índios Kiriri** / Revista do Museu Paulista, vol. 11. - São Paulo, 1948.

SANTOS, Fabricio Lyrio. **Os índios na história da Bahia**. Organizado por Fabrício Lyrio Santos. Cruz das Almas, EDUFRB; Belo Horizonte: Fino Traço, 2016.

SARDINHA, Tony Berber. **Linguística de corpus: histórico e problemática**. D.E.L.T.A., Vol. 16, nº2, 2000. (323-367)

SPANGHERO, Vitória Regina. Produção de dicionários contemporâneos com línguas em contato: o caso do português e das línguas indígenas brasileiras. *In*: OLIVEIRA, Ana Maria Pires de; ISQUERDO, Aparecida Negri. **As ciências do léxico: Lexicologia, Lexicografia, Terminologia**. 5 ed. Campo Grande (MS): Ed. UFMS, 2010.

VILELA, Mário. **Ensino da língua portuguesa: léxico, dicionário, gramática**. Coimbra: Livraria Almedina, 1995.