

AÇÚCAR CANDIL

Tipo de açúcar considerado medicinal, obtido pela cristalização da sacarose. Finalmente ha o assucar mascavado nos engenhos, e resulta do que fica nos fundos dos potes em que se purifica o mais: como pelo contrario o assucar candil é o primeiro na bocca dos mesmos vasos de barro. D. FR. JOÃO DE S. JOSÉ (1869) [1762], VIAGEM E VISITA DO SERTÃO EM O BISPADO DO GRÃO-PARÁ EM 1762 E 1763: ESCRIPTA PELO BISPO D. FR. JOÃO DE S. JOSÉ. [A00_0740 p. 45].

AÇÚCAR DE CARA

FECHADA

Pão de açúcar consistente, o qual não se quebra com facilidade Logo se lhes bota o Affucar por Temperas, como já temos dito; o qual no espaço de três dias endurece diversamente, h mais, outro menos: & ao q´ mais se endurece, & difficultofamente se quebra, chamaõ Affucar de cara fechada; & ao que facilmente com qualquer pancada se quebra, chamaõ Affucar de cara quebrada. ANDRÉ JOÃO ANTONIL (1711) [1711], LIVRO III - CAPITULO I - DAS FORMAS DO AÇÚCAR, E SUA

Aparecida Negri Isquerdo
Elizabete Aparecida Marques
Organizadoras

AS CIÊNCIAS DO LÉXICO

LEXICOLOGIA
LEXICOGRAFIA
TERMINOLOGIA

VOLUME X

Nesta obra, artigos de

Aderlande Pereira Ferraz
Alexandre Melo de Sousa
Álvaro David Hwang
Ana Paula Tribesse Patrício Dargel
Anna Maria Becker Maciel
Aparecida Negri Isquerdo
Ariane Vicente Mota
Celina Márcia de Souza Abbade
Claudia Zavaglia
Cleci Bevilacqua
Clotilde de Almeida Azevedo Murakawa
Elizabete Aparecida Marques
Fábio Henrique de Carvalho Bertonha
Guilherme Fromm
Ieda Maria Alves
Jessica dos Santos Paião
Marcela Moura Torres Paim
Márcia Sipavicius Seide
Márcio Sales Santiago
Maria Cândida Trindade Costa de Seabra
Maria Celeste Augusto
Maria Cristina Parreira da Silva
Maria da Graça Krieger
Maria José Bocorny Finatto
Mariângela de Araújo
Martín Tapia Kwiecien
Odair Luiz Nadin
Pierre-André Buvet
Renato Rodrigues-Pereira
Rosemary Irene Castañeda Zanette
Sabrina Pereira de Abreu
Sara Augusto Carra

CANDIL

açúcar considerado medicinal, obtido pela cristalização do açúcar mascavado. Finalmente ha o assucar mascavado nos engenhos, do que fica nos fundos dos potes em que se e purifica o açúcar, e no pelo contrario o assucar candil é o primeiro na bocca dos vasos de barro. D. FR. JOÃO DE S. JOSÉ (1869) VIAGEM E VISITA DO SERTÃO EM O BISPADO DO PARÁ EM 1762 E 1763; ESCRIPTA PELO BISPO D. JOÃO DE S. JOSÉ. [A00_0740 p. 45].

AÇÚCAR DE CARA FECHADA

o açúcar consistente, o qual não se quebra com facilidade quando lhes bota o Affucar por Temperas, como já temos dito; e no espaço de três dias endurece diverfamente, h mais, e outros menos: & ao que q´mais fe endurece, & difficultosamente fe chamaõ Affucar de cara fechada; & ao que facilmente quer pancada fe quebra, chamaõ Affucar de cara quebrada. ANDRÉ JOÃO ANTONIL (1711) [1711], LIVRO III - CAPITULO I - DAS FORMAS DO AÇÚCAR, E SUA PAIFAGEM DO TENDAL PARA A CAFA DE PURGAR [A00_2578 p. 76].

AÇÚCAR DE CARA QUEBRADA

o açúcar de pouca consistência. Logo fe lhes bota o Affucar por Temperas, como já temos dito; e no qual no espaço de três dias se endurece diverfamente, h mais, outro menos: & ao que q´mais fe endurece, & difficultosamente fe quebra, chamaõ Affucar de cara fechada; & ao que facilmente com qualquer pancada fe quebra, chamaõ Affucar de cara quebrada. ANDRÉ JOÃO ANTONIL (1711), LIVRO III - CAPITULO I - DAS FORMAS DO AÇÚCAR, E SUA PAIFAGEM DO TENDAL PARA A CAFA DE PURGAR [A00_2578 p. 76].

AÇÚCAR ROSADO

o açúcar que é feito a base de rosas, usado como medicamento [...] muito açúcar rosado para purgas, e para não purgar, não têm das outras rosas; cozem as de Alexandria n'agua, e do-lha fora fazem açúcar rosado muito bom com que se cura a febre. PADRE FERNÃO CARDIM (1980) [1585], I – DO QUE SE FAZ NA TERRA DO BRASIL - E DE ALGUMAS COUSAS QUE SE VEIS QUE SE ACHÃO ASSI NA TERRA COMO O AÇÚCAR [A00_0749 p. 59].

AÇÚCAR

o açúcar esfriado no tendal das primitivas casas de engenho, depois de se fazer na casa de purgar, era posto em longas formas dentro das quais se tomava consistência resultando um bolo, donde o seu nome. [...] que o lavrador será obrigado a plantallas de cannas que se fará moer mais do que no Engenho do proprietario que se lhe pertence a metade do assucar que produzirão, além do que he ha da de dar mais daquella metade e com que ficou

AÇÚCAR CANDIL

Tipo de açúcar considerado medicinal, obtido pela cristalização da sacarose. Finalmente ha o assucar mascavado nos engenhos, e resulta do que fica nos fundos dos potes em que se purifica o mais: como pelo contrario o assucar candil é o primeiro na bocca dos mesmos vasos de barro. D. FR. JOÃO DE S. JOSÉ (1869) [1762], VIAGEM E VISITA DO SERTÃO EM O BISPADO DO GRÃO-PARÁ EM 1762 E 1763: ESCRIPTA PELO BISPO D. FR. JOÃO DE S. JOSÉ. [A00_0740 p. 45].

AÇÚCAR DE CARA FECHADA

Pão de açúcar consistente, o qual não se quebra com facilidade Logo se lhes bota o Affucar por Temperas, como já temos dito; o qual no espaço de três dias endurece diversamente, h mais, outro menos: & ao q' mais se endurece, & difficulofsamente se quebra, chamaõ Affucar de cara fechada; & ao que facilmente com qualquer pancada se quebra, chamaõ Affucar de cara quebrada. ANDRÉ JOÃO ANTONIL (1711) [1711], LIVRO III - CAPITULO I - DAS FORMAS DO AÇÚCAR, E SUA

Aparecida Negri Isquierdo
Elizabete Aparecida Marques
Organizadoras

AS CIÊNCIAS DO LÉXICO

LEXICOLOGIA
LEXICOGRAFIA
TERMINOLOGIA

VOLUME X



**UNIVERSIDADE FEDERAL
DE MATO GROSSO DO SUL**

Reitor
Marcelo Augusto Santos Turine

Vice-Reitora
Camila Celeste Brandão Ferreira Ítavo

Obra aprovada pelo
CONSELHO EDITORIAL DA UFMS
Resolução nº 202-COED/AGECOM/UFMS,
de 04 de outubro de 2023.

CONSELHO EDITORIAL
Rose Mara Pinheiro (presidente)
Adriane Angélica Farias Santos Lopes de Queiroz
Ana Rita Coimbra Mota-Castro
Andrés Batista Cheung
Alessandra Regina Borgo
Delasnieve Miranda Daspert de Souza
Elizabete Aparecida Marques
Geraldo Alves Damasceno Junior
Maria Lígia Rodrigues Macedo
William Teixeira

Dados Internacionais de Catalogação na Publicação (CIP) (Coordenadoria de Bibliotecas – UFMS, Campo Grande, MS, Brasil)

As ciências do léxico [recurso eletrônico] : volume X : lexicologia, lexicografia, terminologia / Aparecida Negri Isquierdo, Elizabete Aparecida Marques, organizadoras. – Campo Grande, MS : Ed. UFMS, 2023.

Modo de acesso: <https://repositorio.ufms.br>
Texto em português. Espanhol e francês.
Inclui bibliografia.
ISBN 978-85-7613-637-8

1. Lexicologia. 2. Lexicografia. I. Isquierdo, Aparecida Negri. II. Marques, Elizabete Aparecida

CDD (23) 413.028

Bibliotecária responsável: Tânia Regina de Brito – CRB 1/2.395

Aparecida Negri Isquerdo
Elizabete Aparecida Marques
Organizadoras

AS CIÊNCIAS DO LÉXICO

LEXICOLOGIA
LEXICOGRAFIA
TERMINOLOGIA

VOLUME X

Campo Grande
2023



© das organizadoras
Aparecida Negri Isquierdo
Elizabete Aparecida Marques

Projeto Gráfico, Editoração Eletrônica
Secretaria da Editora UFMS

Revisão
A revisão linguística e ortográfica
é de responsabilidade dos autores

A grafia desta obra foi atualizada conforme o Acordo Ortográfico da Língua Portuguesa, de 1990, que entrou em vigor no Brasil em 1º de janeiro de 2009.

As informações contidas nos textos deste livro são de responsabilidade dos respectivos autores e não expressam a opinião das organizadoras.

Direitos exclusivos para esta edição



Secretaria da Editora UFMS
Av. Costa e Silva, s/nº | Bairro Universitário
Campo Grande - MS, 79070-900
Fone: (67) 3345-7239
e-mail: sedit.agecom@ufms.br

Editora associada à



ISBN: 978-85-7613-637-8
Versão digital: outubro de 2023.

Apoio:



**UNIVERSIDADE FEDERAL
DE MATO GROSSO DO SUL**

Obra contemplada no Edital AGEKOM Nº 01/2021 - Seleção de propostas para apoio a publicação de e-books de Divulgação e Popularização da Ciência pela Editora UFMS.



Este livro está sob a licença Creative Commons, que segue o princípio do acesso público à informação. O livro pode ser compartilhado desde que atribuídos os devidos créditos de autoria. Não é permitida nenhuma forma de alteração ou a sua utilização para fins comerciais. br.creativecommons.org

Dedicamos esta obra, *in memoriam*, às queridas *Maria Teresa Camargo Biderman*, idealizadora da publicação *As Ciências do Léxico*, e *Ana Maria Pinto Pires de Oliveira*, uma das organizadoras do volume 1 da coleção. Para ambas faltou vida para assistirem à expansão da obra! A elas o nosso reconhecimento e respeito!

PREFÁCIO

AS CIÊNCIAS DO LÉXICO: A CONSOLIDAÇÃO DE UMA ÁREA

A publicação do Volume X da coleção *As Ciências do Léxico*, neste ano de 2023, traz consigo um grande número de desafios enfrentados com determinação, coragem e paixão pelos estudos do léxico. A superação dos desafios transformou-se em respostas concretas de um grupo de pesquisadores – professores e estudantes de Pós-Graduação – centrado num campo de investigação multifacetado mas, simultaneamente unitário, pois estabelecido sobre um alicerce comum de pesquisa: o léxico do português do Brasil.

À luz da heterogeneidade constitutiva do componente lexical e dos objetivos particulares de investigação, inscrevem-se as trajetórias das três grandes áreas que, historicamente, dimensionaram o conteúdo desta hoje prestigiosa coleção intitulada *As Ciências do Léxico. Lexicologia, Lexicografia, Terminologia*. Ela foi criada para divulgar os trabalhos apresentados nos encontros bianuais do Grupo de Trabalho de Lexicologia, Lexicografia e Terminologia (GTLEX) da ANPOLL (Associação Nacional de Pós-Graduação em Letras e Linguística).

O conceito de Ciências do Léxico foi concebido pela importante lexicóloga e lexicógrafa Maria Tereza Biderman, que integrou o grupo fundador do GTLEX. Para Maria Tereza, era fundamental enfatizar e divulgar os estudos sobre o componente lexical desenvolvidos na perspectiva da Lexicologia, da Lexicografia e da Terminologia. A própria proponente discorreu na primeira edição do volume 1, datado de 1998, sobre o tripé disciplinar que inaugurou o campo unitário e múltiplo que deu nome à coleção.

Em seus 25 anos de existência, as publicações da obra *As Ciências do Léxico* refletem a expansão de suas fronteiras, ao acolher, cada vez mais, outras subáreas e outras dimensões específicas que se somam ao conhecimento científico do léxico, como a abrangente Onomástica, voltada aos nomes próprios – a Toponímia e a Antroponímia –, com focos nos nomes de lugares e de pessoas, respectivamente; a Neologia, a olhar para as unidades novas e suas diferentes formações morfológicas, típicas do dinamismo lexical e, ainda, o Regionalismo, que retrata muito dos falares regionais.

O enfoque maior dos estudos publicados tem sido dado à contemporaneidade do componente léxico do português brasileiro, embora aspectos históricos dos usos lexicais não sejam esquecidos. A constante observação da natureza, da essência, da constituição formal e da funcionalidade das unidades léxicas responde pela inclusão de suas descrições tanto em sua dimensão simples, quanto sintagmática. Dessa forma, a macroárea das Ciências do Léxico passou também a acolher estudos de Fraseologia.

Todas essas dimensões investigativas, com suas múltiplas variantes, evidenciam as inúmeras possibilidades e focos de análise que o componente léxico das línguas revela e direciona. Descrito em várias perspectivas, os estudos desenvolvidos mantêm o eixo comum de

interesse pelo léxico dito, atestado. Vale dizer, são analisadas unidades léxicas concretamente expressas seja na escrita, seja na oralidade.

Ao contrário do que se possa pensar, não há dispersão nas Ciências do Léxico, mas diversidade produtiva como demonstram os trabalhos apresentados nos encontros do GT de Lexicologia, Lexicografia e Terminologia, fundado em 1986. A riqueza e o avanço investigativo desse Grupo estão escritos desde o lançamento do primeiro volume da Coleção (1998). Destacamos que, em virtude de a obra ter sido contemplada em edital lançado pelo Programa Publicações de Apoio à Formação Inicial e Continuada de Professores do COMPED-INEP/MEC, em 2001, foi publicada, nesse mesmo ano, a segunda edição do primeiro volume da coleção com uma tiragem de 3.000 exemplares. Desse montante, 50% dos exemplares foram comercializados pela Editora UFMS e os 50% restantes distribuídos pelo MEC para bibliotecas de escolas públicas de diferentes níveis de ensino de todo o território nacional.

Ao longo de sua história, a coleção foi se enriquecendo em diferentes aspectos. Inicialmente voltada para a divulgação de pesquisas brasileiras, já desde o volume 2 passou a incorporar trabalhos de relevantes pesquisadores estrangeiros – Mercè Lorente (Espanha), Luis Fernando Lara (México) e André Clas (Canadá) –, que escreveram, respectivamente, sobre as três subáreas clássicas que inicialmente articulavam as “Ciências do Léxico”. A partir desse número, a participação de pesquisadores estrangeiros, convidados pelos membros do GT, tem sido constante e enriquecedora.

No contexto brasileiro, jovens mestres e doutores formados pelos primeiros e mais experientes membros do GTLEX foram também se integrando ao grupo, constituindo novas gerações de pesquisadores do léxico que, sucessivamente, vêm se agregando ao GTLEX,

ampliando, assim, o conjunto dos autores da coleção. Como se observa, é uma nova geração atuante, inovadora em propostas e que mantém viva a tradição e o interesse pelo conhecimento do léxico.

A dinamicidade da coleção é também observada nas temáticas abordadas. Sempre aberta a novas realidades, a coleção reflete tendências contemporâneas e integra novas perspectivas dos estudos lexicais, a exemplo dos trabalhos terminológicos relacionados às línguas de sinais, os neologismos observados nas redes sociais, a incorporação da multimodalidade pela lexicografia.

De igual modo, e já há algum tempo, as Ciências do Léxico vêm abrindo diálogos importantes como fez com a Linguística de Corpus, introduzindo condições de aplicação para ferramentas lexicais, e também com a Tradução e com teorias de texto e discurso. Neste último caso, trouxe a perspectiva da Terminologia Textual para contribuir metodologicamente com trabalhos de elaboração de repertórios de termos e fraseologias presentes nas linguagens especializadas. Assim, também o próprio estudo das linguagens especializadas redimensionou o conhecimento do léxico.

Entre vários outros horizontes de trabalhos, e novos diálogos com diferentes áreas da linguística como a semântica, já que léxico e significado se materializam conjuntamente, desde o início das publicações, encontra-se um tema que percorre com assiduidade os vários volumes da coleção. Trata-se de preocupações e proposições pedagógicas que envolvem o estudo do léxico tanto no plano da língua materna, quanto estrangeira. A importância da dimensão pedagógica dos estudos lexicais permeia a lexicologia, mas se sobressai na lexicografia, nas descrições e proposições de uso de dicionários adequados para o ensino. Nesse contexto, os trabalhos do GTLEX foram capazes, inclusive, de influenciar políticas públicas determinantes da área

de dicionários do Programa Nacional do Livro Didático (PNLD) do Ministério da Educação.

Em síntese, a publicação deste volume X nos permite desenhar o sólido panorama construído pelas Ciências do Léxico e sua consolidação no âmbito da Linguística Aplicada do Brasil. Aqui não se trata de autoelogio, mas da qualidade e regularidade de uma publicação que, em seus 24 anos, não apenas manteve a regularidade de publicação, como tornou-se referência nos e para os estudos lexicais do país, alcançando prestígio nacional. No âmbito da CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), nas versões preliminares do sistema de classificação de livros (Qualis Livros), a obra recebeu avaliações muito positivas. A Editora UFMS tem buscado cumprir, nos últimos volumes, requisitos adotados pela CAPES para avaliação desse tipo de produção científica e, no momento, aguarda com expectativa o documento final do Qualis Livros que será o parâmetro para a avaliação desse tipo de produção no âmbito dos Programas de pós-graduação *stricto sensu* no Brasil.

Todo esse prestígio nacional alcançado pelas Ciências do Léxico deve-se fortemente a alguns nomes que não podem deixar de ser lembrados. Entre eles, a maior referência é Aparecida Negri Isquerdo. Como “editora-chefe”, é a grande responsável pela regularidade e qualidade das publicações. Desde o primeiro volume, Aparecida contou com o importante apoio da Editora da Universidade Federal de Mato Grosso do Sul (Campo Grande). A concretização dos volumes da coleção recebeu a colaboração de Ana Maria Pinto Pires de Oliveira, em seu primeiro volume, e, nos volumes subsequentes, a organização passou a ser de responsabilidade da Aparecida, conjuntamente, com os sucessivos coordenadores do GTLEX.

Neste momento, no marco da publicação do Volume X, pode-se afirmar sem contestações que a coleção *As Ciências do Léxico* é uma referência bibliográfica essencial para os estudiosos do Léxico e é também o testemunho da consolidação da área, bem como das positivas relações dos membros, advindos das diversas universidades brasileiras, e que integram e fortalecem o GT de Lexicologia, Lexicografia e Terminologia da ANPOLL.

Ieda Maria Alves
Maria da Graça Krieger

SUMÁRIO

Apresentação..... 15

PARTE I – LEXICOLOGIA

Entretien avec un robot

Pierre-André Buvet..... 28

Analogia e desautomatização: processos neológicos convergentes

Aderlande Pereira Ferraz..... 56

Toponímia em Libras: iconicidade e empréstimos na formação dos sinais

Alexandre Melo de Sousa 80

O nome próprio de pessoa na toponímia urbana: interfaces entre processo denominativo e ideologia

Ana Paula Tribesse Patrício Dargel; Aparecida Negri Isquerdo..... 102

O ATOBAH e as nomeações genéricas na hidronímia baiana

Celina Márcia de Souza Abbade 124

Fraseologismos no *corpus* do projeto Atlas linguístico do Brasil

Marcela Moura Torres Paim..... 143

Nomeações secundárias em contextos migratórios e sua representação literária

Márcia Sipavicius Seide 159

Hidrotoponímia: estudo de casos

Maria Cândida Trindade Costa de Seabra 178

Experiência do ensino remoto de lexicologia e lexicografia

Rosemary Irene Castañeda Zanette..... 205

PARTE II – LEXICOGRAFIA

Obras de referência portuguesas no *Hobson-Jobson*: uma abordagem lexicográfica

Maria Celeste Augusto 232

Dicionário monolíngue de língua estrangeira: a *face visível* e a *face invisível* das palavras

Álvaro David Hwang; Maria Cristina Parreira 254

***Gíria*: marca de uso presente em dicionários escolares**

Claudia Zavaglia; Fábio Henrique de Carvalho Bertonha..... 277

Memórias da “viagem” da construção do Dicionário Histórico do Português do Brasil (DHPB) e suas repercussões lexicográficas

Clotilde de A. Azevedo Murakawa..... 304

Expressões idiomáticas em dicionários: questões iniciais em direção à fraseografia pedagógica

Elizabete Aparecida Marques; Jessica dos Santos Paião 324

El enfoque sincrónico para el establecimiento de la homonimia destinada al *DMHE*

Renato Rodrigues-Pereira; Odair Luiz Nadin;

Martín Tapia Kwiecien 347

PARTE III – TERMINOLOGIA

A linguagem do patrimônio cultural de bens móveis em papel: construção de uma base de dados terminológica

Anna Maria Becker Maciel; Cleci Regina Bevilacqua..... 375

MetaLex: por uma metaferramenta lexical

Guilherme Fromm 399

A linguagem em tempos de pandemia: registro da terminologia do novo coronavírus e da COVID-19

Márcio Sales Santiago 417

Desde quando medicamento não é remédio?

Estudo computacional sobre o vocabulário de textos médicos em português do século XVIII

Maria José Bocorny Finatto 438

O sangue é um líquido ou um tecido? A elaboração de definições para um dicionário terminológico das ciências naturais para professores

Mariângela de Araújo; Ariane Vicente Mota 486

Frasemas semântico-lexicais composicionais no âmbito da teoria sentido-texto: estudo de alguns clichês do léxico do português brasileiro

Sabrina Pereira de Abreu; Sara Augusto Carra 504

Os autores..... 522

APRESENTAÇÃO

Decorridos vinte e cinco anos da publicação da primeira edição do livro que acabou se configurando como o volume 1 da coleção *As Ciências do Léxico. Lexicologia, Lexicografia e Terminologia*, vem a lume o volume X que, a exemplo dos demais, tem a palavra como foco principal de abordagem nos 21 artigos que compõem a obra. Nesse sentido, em sintonia com os objetivos da coleção, o volume X aborda a *palavra* sob diferentes enfoques teóricos.

A coleção *As Ciências do Léxico* chega, assim, ao seu décimo volume mantendo-se fiel aos objetivos e à estrutura da primeira edição (1998), à época definidos pela sua idealizadora, Maria Tereza Camargo Biderman, em função do propósito da publicação. A saudosa pesquisadora do léxico confiou a organização da obra a duas jovens doutoras, com o argumento de que essa tarefa seria para elas importante em termos de currículo e, para a Editora da UFMS, uma instituição que buscava o seu espaço no meio acadêmico, a publicação contribuiria para a projeção da editora no cenário nacional. A proposta inicial era a organização de uma obra única, conforme idealizara Maria Tereza, então coordenadora do GTLEX – Grupo de Trabalho em Lexicologia, Lexicografia e Terminologia da ANPOLL – Associação

Nacional de Pesquisa e Pós-Graduação em Letras e Linguística que, por sua vez, congrega docentes permanentes em Programas de Pós-Graduação em Letras e Linguística vinculados a universidades das cinco regiões do Brasil.

Todavia, quis a história que os caminhos fossem outros: a obra, com a chancela da Editora UFMS e mediante o compromisso dos membros do GTLEX, foi transformada em coleção e chega este ano ao seu décimo volume. Maria Tereza e Ana Maria (uma das organizadoras do primeiro volume) já vivem no plano espiritual e devem estar comemorando o sucesso de uma coleção que acabou se transformando em veículo de difusão de pesquisas lexicais no Brasil, uma referência para os pesquisadores interessados em investigações sobre a palavra em diferentes perspectivas. Os nove volumes já publicados disseminaram um montante de 241 trabalhos científicos, 94 (39%) que versam sobre temas relacionados à subárea Lexicologia que também abriga a Fraseologia e a Onomástica; 60 (24,9%) que tiveram como foco o léxico em estado de dicionário, portanto, vinculados à subárea Lexicografia, e 87 (36,1%) que têm como foco o vocabulário especializado de diferentes áreas do saber. Desse montante de 241 artigos, 33 (13,6%) foram produzidos por pesquisadores estrangeiros, em sua maioria europeus, que participaram das publicações como convidados a partir do volume 2, em média, pelo menos um de cada subárea contemplada pela coleção.

A palavra aparece nesta publicação, assim como nas anteriores, discutida em suas diferentes dimensões e funcionalidades, o que pode ser observado nas abordagens contempladas pelos trabalhos que compõem este volume. Obedecendo ao plano da obra, os textos estão distribuídos segundo uma das três grandes partes em torno das quais os textos são organizados, disponibilizando, assim, ao público-alvo, na Parte I – Lexicologia – nove artigos que versam sobre neologismos,

fraseologia, toponímia e antroponímia, ensino do léxico e competência lexical no âmbito da inteligência artificial; na Parte II – Lexicografia – seis trabalhos de focalizam a obra lexicográfica a partir de diferentes olhares, o mesmo ocorrendo com a Parte III – Terminologia – a que foram reunidos seis artigos que abordam o léxico de especialidade em diferentes domínios.

Desses trabalhos, dois são de autoria de pesquisadores estrangeiros, um francês e uma portuguesa radicada na Holanda, além de um pesquisador argentino que escreveu um trabalho em coautoria com dois estudiosos brasileiros. Os trabalhos de pesquisadores brasileiros que integram este volume, a exemplo dos anteriores, são produtos de pesquisas desenvolvidas e discutidas no Grupo de Pesquisa de Lexicologia, Lexicografia e Terminologia da ANPOLL.

Abrindo a primeira parte, o texto **Entretien avec un robot**, de autoria de Buvet, apresenta um sistema de diálogo implementado em um robô humanoide projetado para ajudar idosos que vivem em ambiente semi-hospitalar. Em um primeiro momento, o autor especifica como o sistema de diálogo está posicionado em relação a outras interfaces homem-máquina para, em um segundo momento, discutir a análise de fatos linguísticos que permitiram modelar uma conversa entre um ser humano e uma máquina, com especial atenção para o nível lexical. Em terceiro lugar, o texto destaca os principais aspectos tecnológicos do sistema.

Na sequência, em **Analogia e desautomatização: processos neológicos convergentes**, Ferraz discute a neologia formal manifestada em textos publicitários na mídia eletrônica (sítios e redes sociais na internet). Criações que decorrem da alteração de uma forma preexistente, por meio do espelhamento lexical, como os casos de analogia e desautomatização, foram analisadas, numa perspectiva voltada para

o desenvolvimento da competência lexical. Desse modo, o trabalho circunscreve-se à análise desses dois processos neológicos que, embora muito raramente tratados pelas gramáticas tradicionais, são produtivos no português brasileiro contemporâneo, como se verificam em textos publicitários da mídia eletrônica, que constituíram o *corpus* de análise. Em síntese, o escopo do trabalho é caracterizar, no âmbito da inovação lexical, a principal relação entre analogia e desautomatização: a convergência de procedimentos na gênese dos produtos neológicos gerados.

Na sequência, o texto **Toponímia em Libras: iconicidade e empréstimos na formação dos sinais**, de autoria de Sousa, focaliza a questão do mapeamento icônico e sua relação com a estrutura morfofonológica dos topônimos em Libras, valendo-se, para tanto, de exemplos de sinais toponímicos utilizados por surdos nativos que nomeiam espaços acreanos, provenientes do Projeto Toponímia em Libras. Com base nos estudos de Taub (2001), Perniss (2007) e Quadros (2019), o trabalho discute o processo de formação dos itens lexicais icônicos e da relação entre a iconicidade e as taxionomias propostas por Dick (1990; 1992) para a classificação de topônimos.

Abordando também a temática da toponímia, o texto de Dargel e Isquerdo **O nome próprio de pessoa na toponímia urbana: interfaces entre processo denominativo e ideologia** centra-se no exame de dados da toponímia urbana de cidades sul-mato-grossenses e tem como propósito analisar tendências evidenciadas por formações antropotoponímicas na toponímia urbana, considerando duas taxes toponímicas, a dos antropotopônimos e a dos axiotopônimos (Dick, 1990; 1992). Para tanto, examinam os dados a partir de três eixos: i) produtividade; ii) a *projeção* da pessoa homenageada pela toponímia em nível local, regional, nacional, internacional e, iii) a *categoria* profissional das personalidades homenageadas.

Ainda no âmbito dos estudos toponímicos, Abadde, no texto **O ATOBAH e as nomeações genéricas na hidronímia baiana**, aborda as nomeações dos principais termos genéricos existentes na hidronímia baiana. Considerando a riqueza e a variação do sistema hidrográfico do Estado, a autora destaca que a população acaba generalizando os termos que nomeiam correntes hídricas, muitas vezes não fazendo distinção entre acidentes geográficos nomeados como rio, riacho ou córrego, por exemplo. Nesse cenário, o trabalho discute termos catalogados no *corpus* do Projeto ATOBAH (Atlas Toponímico da Bahia) que se configuram como nomes geográficos e que, no trabalho, foram organizados pela autora em verbetes, cujas definições foram elaboradas a partir de informações fornecidas por Cunha (2010), Houaiss e Villar (2001), IBGE (2018), Souza 2004 [1939].

Na sequência, o texto **Fraseologismos no corpus do Projeto Atlas Linguístico do Brasil**, de autoria de Paim, tem foco na variação lexical e, buscando relações entre Fraseologia e Dialetoлогия, analisa fraseologismos identificados na base nos dados do Projeto ALiB – Atlas Linguístico do Brasil, área temática “ciclos da vida”, mais especificamente, unidades fraseológicas utilizadas como denominação para “menstruação”, “menopausa” e “filho caçula”, documentadas em 25 capitais do Brasil, com brasileiros estratificados por sexo, faixa etária e nível de escolaridade.

Já no texto **Nomeações secundárias em contextos migratórios e sua representação literária**, Seide analisa essa categoria de nomeações e sua representação literária no contexto de migração dos lituanos ao Brasil e a países de língua inglesa. Para tanto, a autora analisou as seguintes categorias de nomes próprios de pessoas: nome civil (formado por prenome e nome de família), hipocorístico, apelido e pseudônimo, concluindo que, dependendo da política linguística de

facto adotada em cada país, as mudanças oficiais de nomes são possíveis, impossíveis ou obrigatórias.

Na continuidade, Seabra, no texto **Hidronímia: estudo de casos**, apresenta um estudo sobre nomes de lugares motivados pela água (hidrotopônimos) na toponímia mineira com base em denominações toponímicas presentes no banco de dados do Projeto ATEMIG (Atlas Toponímico do Estado de Minas Gerais), nomes esses que sintetizam a percepção do denominador em épocas pretéritas. Para tanto, a autora vale-se da metodologia dos “estudos de caso” que permitem a observação de detalhes da trajetória percorrida por um topônimo ao longo do tempo.

Fechando a Parte I da obra, o texto **Experiência do ensino remoto de Lexicologia e Lexicografia**, assinado por Zanette, apresenta e discute reflexões acerca de experiência de ensino de língua inglesa na modalidade remota, no período da Pandemia Covid-19, experiência apoiada na legislação nacional e local, bem como em teóricos sobre o ensino emergencial remoto, aliado ao uso das tecnologias digitais da informação e da comunicação (TDIC). Para tanto, a autora descreve e analisa os três momentos da atividade (preparação; desenvolvimento e avaliação da disciplina ministrada) com foco nas reflexões teóricas que sustentaram a disciplina e respectivas descrição e reflexão das ações desenvolvidas.

Iniciando a Parte II da obra, que agrupa os trabalhos relacionados à área da Lexicografia, o texto **Obras de referência portuguesas no Hobson-Jobson: uma abordagem lexicográfica**, produzido por Augusto, examina os dois títulos portugueses mais mencionados no *Glossary: Colóquios dos Simples e Drogas* de Garcia de Orta e *Décadas da Asia* de João de Barros, com o objetivo primordial de refletir sobre o modo como os autores do *Glossary* se serviram dos referidos textos

portugueses e, ao mesmo tempo, procurar justificativas para o epíteto “alambique” dado pelos dois lexicógrafos ingleses à Língua Portuguesa, mais precisamente ao Indo-Português. Destaca ainda a autora a função de veículo transmissor do Português, ao servir de intermediário entre as línguas “exóticas” e as línguas europeias, o que, salvo raríssimas exceções, não tem recebido a atenção que merece na maior parte dos estudos linguísticos sobre essa temática.

O segundo artigo, **Dicionário monolíngue de língua estrangeira: a face visível e a face invisível das palavras**, assinado por Hwang e Parreira, trata de questões de possíveis interesses do usuário que procura encontrar no dicionário monolíngue de língua estrangeira uma ferramenta transformadora de sua experiência com a língua descrita, com base, sobretudo, nos aspectos da descrição que revelam o potencial da obra lexicográfica para o desenvolvimento de sua capacidade de produção na língua estrangeira. Os aspectos relativos ao que os autores chamam de *face visível* e *face invisível* das palavras têm relação com o controle das informações, aspecto do fazer lexicográfico que tem um papel importante na coesão interna do discurso do dicionário.

Em **Gíria: marcas de uso presente em dicionários escolares**, terceiro trabalho da segunda seção da obra, Zavaglia e Bertonha discorrem sobre a questão das marcas de uso nos dicionários escolares, mais especificamente a ‘gíria’, de natureza diastrática. Para tanto, selecionam nove unidades lexicográficas (abobrinha, bico, biruta, duro, irado, morar, pau, pronto e zoar) com a marca de uso ‘gíria’ em oito dicionários monolíngues: Geiger (2011), Ferreira (2008; 2010), Braga e Magalhães (2011), Biderman (2009), Houaiss (2009; 2021) e Aulete (2021). Com base em literatura especializada nessa temática, os autores analisam, pois, a utilização da marca de uso ‘gíria’, buscando verificar em que proporção ela contribui para o entendimento e restrição contextual.

Por sua vez, o texto **Memórias da “viagem” da construção do Dicionário Histórico do Português do Brasil (DHPB) e suas repercussões lexicográficas**, produzido por Murakawa, traz uma narrativa de caráter memorialista sobre a história do DHPB – Dicionário Histórico do Português do Brasil (Biderman; Murakawa, 2021), em que a autora resgata o itinerário da “viagem”, metáfora utilizada para descrever as diferentes e inquietantes fases que marcaram a produção da obra, desde a formatação do projeto, realizada por Biderman em 2005, até a publicação da obra em rede, conquistada por Murakawa e pesquisadores da UNESP em 2021.

Na sequência, Marques e Paião, com base nos pressupostos da Lexicografia Pedagógica e da Fraseografia, no trabalho **Expressões idiomáticas em dicionários: questões iniciais em direção à Fraseografia Pedagógica**, analisam o tratamento de expressões idiomáticas em dicionários escolares, tanto da língua portuguesa como da língua espanhola. O estudo mostra que as unidades fraseológicas, especificamente as EI, não recebem um tratamento lexicográfico sistematizado e adequado para os aprendizes de língua, seja estrangeira (espanhol) ou materna, evidenciando que o registro dessas unidades em dicionários pedagógicos exige novas práticas, que conjuguem pressupostos teóricos-metodológicos de diversos saberes.

Encerrando a Parte II, no texto **El enfoque sincrónico para el establecimiento de la homonimia destinada al DMHE**, Rodrigues-Pereira, Nadin e Kwiecien discutem resultados da segunda etapa de uma pesquisa empírica realizada com o objetivo de demonstrar a possibilidade de se considerar o critério sincrónico para a definição de *Unidades Léxicas Homônimas* (ULH), no processo de inventário de candidatos a homônimos que possam vir a compor a nomenclatura de dicionários pedagógicos do espanhol como língua

estrangeira, em especial o DMHE – *Diccionario monolingüe de formas homónimas en español para aprendices brasileños*, em elaboração.

Iniciando a Parte III desta coletânea, o trabalho **A linguagem do patrimônio cultural de bens móveis em papel: construção de uma base de dados terminológica**, produzido por Maciel e Bevilacqua, discute as bases teóricas que sustentam os projetos *A linguagem do patrimônio cultural brasileiro: conservação dos bens culturais móveis* (2016-2019) e *Conservação dos bens culturais móveis em papel: identificação e representação de unidades fraseológicas especializadas* (UFES) (2019-2021), considerando as etapas metodológicas da pesquisa, do tratamento dos dados e os resultados alcançados que culminaram na construção da base de dados. As autoras dão destaque para questões teórico-metodológicas que implicam na elaboração de um produto terminográfico multilíngue e *online* contendo unidades especializadas de uma área em consolidação e que ainda requer reconhecimento no Brasil: a Conservação e Restauração de Bens Culturais.

Por sua vez, o texto **MetaLex: por uma metaferramenta lexical**, assinado por Fromm, apresenta e discute a proposta teórica de elaboração de uma metaferramenta lexical, MetaLex, baseada ou não em *corpora*, que possa fornecer aos especialistas das áreas de Ciências do Léxico a possibilidade de produzirem obras lexicográficas (dicionários), terminográficas (vocabulários) e repositórios para trabalhos onomásticos. No desenvolvimento do trabalho, o autor discute e delimita as noções de recursos, ferramentas e ambientes, disponíveis *online* ou para *download* para a classificação de material já disponível. Para a proposição final da MetaLex, o autor desenvolve uma primeira proposta de fluxograma voltado para a Lexicografia, com o objetivo de sistematizar propostas para o desenvolvimento computacional da metaferramenta proposta.

No texto **A linguagem em tempos de pandemia: registro da terminologia do novo coronavírus e da COVID-19**, Santiago apresenta e discute a ferramenta elaborada por ele para o registro dos termos relacionados à pandemia do COVID-19, a ficha terminológica, dada a sua função de armazenar as informações que posteriormente podem ser utilizadas para a elaboração do glossário. No caso do trabalho em pauta, de termos médicos que circularam tanto entre a comunidade científica quanto na sociedade como um todo em decorrência da referida pandemia, despertando interesse e curiosidade na população, o que explica a funcionalidade denominativa e conceitual exercida pelos termos das diferentes áreas de especialidades e que justifica o fato de o mundo globalizado ter conferido relevância às terminologias. O autor destaca também as informações linguísticas e documentais a serem contempladas nos verbetes do glossário em processo de produção.

Na sequência, apoiando-se na Terminologia Diacrônica, Finatto, no texto **Desde quando medicamento não é remédio? Estudo computacional sobre o vocabulário de textos médicos em português do século XVIII**, relata um conjunto de experimentos de extração automática de informação, com técnicas de Processamento de Linguagem Natural (PLN), a partir de uma obra de Medicina de 1707, em português, com mais de 600 páginas, escrita pelo médico Curvo Semedo (1635-1719). Com base nesse *corpus*, a autora discute o emprego dos termos *medicamento* e *remédio* em contraponto com obras coevas, *corpora* textuais e bases lexicográficas dedicados ao português do século XVIII e demonstra uma série de elementos sintático-semânticos que podem auxiliar a caracterizar a complexidade desse tipo de texto e discurso comparados a outras obras especializadas tanto do período quanto da atualidade.

No trabalho **O sangue é um líquido ou um tecido? A elaboração de definições para um dicionário terminológico das ciências naturais para professores**, Araújo e Mota discutem resultados preliminares do projeto de elaboração de um dicionário terminológico das Ciências Naturais tendo como público-alvo professores do ensino fundamental. As autoras focalizam a metodologia utilizada na elaboração das definições, alguns mapas conceituais elaborados e exemplos de definições, além de pontuarem as vantagens da metodologia adotada, sem deixar de apontar os problemas e os desafios encontrados na execução da pesquisa.

Encerrando a obra, com base nos pressupostos da Teoria Sentido-Texto (TST), Abreu e Carra, no texto **Frasemas semântico-lexicais posicionais no âmbito da teoria sentido-texto: estudo de alguns clichês do léxico do português brasileiro**, discutem frasemas inventariados como “lugares-comuns” ou “frases feitas” em um dicionário organizado por Werneck (2014) durante mais de três décadas. As autoras ponderam que frases e segmentos são de difícil classificação no âmbito dos estudos lexicais, tanto em função da extensão diversificada que apresentam (podem ser frases inteiras, um sintagma, uma única palavra etc.) quanto em função do tipo de sentido que veiculam, por não serem opacos, razão pela qual o seu sentido só se atualiza em contextos pragmáticos específicos.

Neste volume da coleção *As Ciências do Léxico*, em especial, rendemos homenagem às pesquisadoras que têm contribuído de forma particular para a escrita da história da publicação, por meio de parcerias com *Aparecida Negri Isquardo/UFMS* na organização do livro a partir do volume II. Por decisão dos membros efetivos do GTLEX, em assembleia, com o falecimento da Ana Maria Pinto Pires de Oliveira, os coordenadores do GT de cada biênio passaram a atuar na organização da publicação oficial do grupo, inclusive tendo solidificado, em

alguns dos volumes, parcerias institucionais com a Editora da UFMS: *Maria da Graça Krieger/UFRGS* (vol. II); *Ieda Maria Alves/USP* (vol. III); *Maria José Bocorny Finatto/UFRGS* (vol. IV); *Lídia Maria Barros/UNESP* (vol. V); *Maria Cândida Trindade Costa de Seabra/UFMG* (vol. VI); *Giselle Olivia Mantovani Dal Corno/UCS* (vol. VII e VIII); *Celina Márcia de Souza Abadde/UNEB* (vol. IX) e *Elizabete da Silva Marques/UFMS* (vol. X). A todas as parceiras de sonhos e de trabalho o nosso agradecimento, respeito e gratidão!

Estende-se, também, um especial agradecimento das organizadoras deste volume aos autores que, em meio a tantos percalços pós-pandemia, se dispuseram a participar desta publicação, partilhando os seus saberes nas áreas da Lexicologia, Lexicografia e Terminologia!

Um agradecimento especial também à Editora UFMS que, há mais de duas décadas, vem mantendo o compromisso assumido em 2004, quando da publicação do vol. 2 da coleção, com os membros do GTLEX no sentido de garantir a publicação regular de uma obra de referência que chega ao número dez com público garantido. Os dois últimos volumes da coleção *As Ciências do Léxico* foram publicados no formato e-book com recursos do Edital PUBLICA UFMS.

Por fim, registramos a nossa especial satisfação por intermediar a publicação do volume X da nossa coleção e desejamos que o seu conteúdo atenda as expectativas do público-alvo a que se destina: pesquisadores das diferentes áreas das Ciências do Léxico, estudantes de graduação e de pós-graduação e demais interessados no estudo da palavra em suas diferentes magias e abordagens!

Aparecida Negri Isquerdo
Elizabete Aparecida Marques
Organizadoras

PAFFAGEM DO
TENDAL PARA A
CAFA DE PURGAR
[A00_2578 p. 76].

AÇÚCAR DE CARA QUEBRADA

Pão de açúcar de pouca consistência. Logo se lhes bota o açúcar por Temperas, como já temos dito; o qual no espaço de três dias endurece diversamente, h mais, outro menos: & ao q' mais se endurece, & difficulosamente se quebra, chamaõ **Açúcar de cara fechada**; & ao que facilmente com qualquer pancada se quebra, chamaõ Açúcar de cara quebrada. ANDRÉ JOÃO ANTONIL (1711) [1711], LIVRO III - CAPITULO I - DAS FORMAS DO AÇÚCAR, E SUA PAFFAGEM DO TENDAL PARA A CAFA DE PURGAR [A00_2578 p. 76].

AÇÚCAR ROSADO

Açúcar que é feito a base de rosas, usado como medicamento [...] e fazem muito **açúcar rosado** para purgas, e para não purgar, porque não têm das outras rosas; cozem as de Alexandria n'agua, e botando-lha fora fazem **açúcar rosado** muito bom com que não purgão. PADRE FERNÃO CARDIM

PARTE I LEXICOLOGIA

ENTRETIEN AVEC UN ROBOT

Pierre-André Buvet

INTRODUCTION

Simuler la parole humaine, à défaut de la reproduire, est un enjeu majeur pour la robotique. Lorsqu'ils seront effectivement des compagnons de vie, les robots humanoïdes communiqueront oralement avec des êtres humains. Dans cette optique, nous développons, de façon expérimentale, un système de dialogue en français qui permet, d'une part, de simuler la capacité humaine du langage en termes de réception et d'émission de messages, et, d'autre part, d'avoir des interactions verbales élaborées avec leurs usagers. A terme, ce système sera implémenté dans des robots humanoïdes destinés à accompagner des personnes âgées dans leur vie quotidienne.

Pour atteindre cet objectif, nous nous appuyons sur la compréhension du langage humain. Les études en linguistique mettent en avant des oppositions comme langue *vs* parole, préconstruit *vs* construit ou compétence *vs* performance. Ces oppositions servent à présenter le langage comme un système de communication fondé sur un nombre fini de règles à partir desquelles sont produits des messages en nombre infini. D'une façon universelle ou au sein d'une langue donnée, les règles sont identifiées, décrites et expliquées dans différents modèles. Les modèles s'accordent sur

l'existence d'invariants qui structurent le langage, mais différent en partie sur leur nature et leur mode de fonctionnement. Nous utilisons ici un modèle s'inspirant des travaux de Zellig Z. Harris (Harris, 1976), Maurice Gross (Gross, 1981) et des récents travaux de Robert Martin (Martin, 2016). A partir de ce modèle, sont construites des représentations formalisées des faits de langue. Elles sont exploitées par des outils informatiques développés dans le cadre du traitement de l'information non structurée (Silberstein, 2016).

Nous présentons ici la version 3 de notre système de dialogue. Sa composante conversationnelle la distingue de la version 2. La nouvelle version, à l'instar de la précédente, associe un message entrant, formulé par l'utilisateur, à un message sortant, formulé par la machine. Dans ce cadre, le système est capable de satisfaire oralement à une demande d'information. Désormais, il permet d'enchaîner les messages oraux sous une forme conversationnelle. Ce qui inclut, en outre, des messages relatifs à une demande d'information, les messages relatifs à une demande d'action ainsi que les messages propres au genre de la discussion. De plus, aussi bien l'utilisateur que la machine ont la possibilité d'initier la conversation. Après avoir précisé comment nous nous positionnons d'un point de vue méthodologique, nous présentons le modèle de données utilisé pour développer le système de dialogue. Nous expliquons ensuite comment fonctionne ce système.

1. POSITIONNEMENT MÉTHODOLOGIQUE

Pour doter un robot de la parole, cinq compétences langagières sont identifiées: la compétence lexicale, la compétence grammaticale, la compétence sémantique, la compétence pragmatique et la compétence dialogique. Ces compétences interviennent pour

échanger des informations sous la forme de messages écrits ou oraux. De tels messages reposent *a minima* sur l'intervention de deux acteurs: le locuteur, et l'allocutaire. Le premier émet le message en y encodant les informations qu'il souhaite transmettre, le second reçoit le message en le décodant afin d'identifier les informations qui lui sont transmises¹. Le locuteur et l'allocutaire sont deux interlocuteurs tels que leurs rôles sont interchangeables, le locuteur devenant l'allocutaire et l'allocutaire se transformant en locuteur. C'est notamment le cas lorsque la communication prend la forme d'un dialogue, c'est-à-dire lorsque les deux interlocuteurs interagissent verbalement entre eux. L'interaction se manifeste par une succession de messages qui sont émis à tour de rôle par les interlocuteurs. Quatre mécanismes sont à l'œuvre pendant un échange d'informations: 1) l'interprétation des messages entrants; 2) la formulation des messages sortants; 3) la mise en relation des messages entrants avec les messages sortants; 4) la coordination du flux conversationnel. Les mécanismes d'interprétation et de formulation exploitent l'ensemble des cinq compétences langagières mentionnées ci-dessus, les mécanismes de mise en relation et de coordination en exploitent au moins une partie.

Les cinq compétences sont interdépendantes dans la mesure où aucune ne peut fonctionner en dehors des quatre autres. Cette interdépendance contribue à expliquer la complexité langagière (Mufwene, 2010). Les études des faits de langues sous-estiment généralement l'imbrication de ces compétences en les représentant de façon séparée du point de vue de leurs propriétés et de celui de leur mode fonctionnement. On le constate notamment en didactique des langues et dans les approches analytiques de la linguistique théorique. Elles le sont également dans les travaux en traitement automatique

¹ Sur le schéma de la communication, cf. https://fr.wikipedia.org/wiki/Sch%C3%A9ma_de_Jakobson.

des langues qui analysent les faits de langue selon différents niveaux: le niveau morphologique, le niveau syntaxique, le niveau sémantique et le niveau pragmatique (Pierrel, 2000).

La compétence lexicale est un savoir linguistique qui consiste, d'une part, à acquérir le vocabulaire d'une langue et, d'autre part, à l'employer en discours de façon appropriée. La compétence grammaticale est un autre savoir linguistique qui comporte également deux volets: l'acquisition des règles morphosyntaxiques qui régissent la formation des énoncés et la maîtrise de ces règles pour formuler des énoncés bien formés (Parisse, 2009). Du seul point de vue de ces deux compétences, l'interdépendance est manifeste². Il y a croisement des savoirs parce qu'employer un mot à bon escient implique de savoir dans quelles conditions grammaticales on peut le faire. Par exemple, en français, une acception de *pleuvoir* est liée à une tournure impersonnelle (*il pleut sur Nantes*) tandis qu'une autre acception de ce verbe est rattachée à une construction à verbe indirectement transitif telle que le sujet est obligatoirement un groupe nominal (*les coups pleuvent sur le pauvre Bill*). Inversement, maîtriser une règle morphosyntaxique nécessite de connaître son champ d'application lexicale. Par exemple, en français, la tournure passive ne concerne pas tous les verbes de la même façon. Ainsi, un des emplois du verbe *voler* est facilement passivable tandis que le verbe synonyme *subtiliser* l'est plus difficilement.

La compétence sémantique a de nombreux aspects, car elle intéresse toutes les dimensions de la langue: la morphologie, le lexique, la syntaxe, la phraséologie, le texte, le dialogue, ... L'interdépendance entre compétence lexicale, compétence grammaticale et compétence sémantique se manifeste, entre autres, dans l'opposition entre

² Les travaux sur le lexique grammaire sur le lexique-grammaire illustrent cette analyse, cf. <https://fr.wikipedia.org/wiki/Lexique-grammaire>.

signification lexicale (celle du verbe *pleuvoir*, en tant que phénomène météorologique, dans *Il pleut sur Nantes*) et signification grammaticale (celle du verbe *pleuvoir*, en tant que marqueur de l'aspect fréquentatif, dans *Les coups pleuvent sur le pauvre Bill*) (Blanco, 2009).

La compétence pragmatique concerne l'intégration de ce qui est extralinguistique dans la production langagière et réciproquement. Dans le premier cas de figure, il s'agit de la prise en compte de la situation de communication dans l'analyse d'un message (la séquence *j'ai mal à la tête* s'interprète différemment lorsqu'elle est prononcée après une longue séquence de travail ou bien dans une pharmacie) (Bracops, 2010). Dans le second cas de figure, il est question de l'influence d'un message sur une situation de communication (une déclaration officielle de mariage le rend effectif) (Austin, 1962). Par définition, la compétence pragmatique est interdépendante des autres compétences. La compétence dialogique est un savoir linguistique qui permet à un être de langage d'interagir avec un autre. Elle incorpore les quatre autres compétences, et, par voie de conséquence, interagit avec elles, tout en présentant des caractéristiques qui lui sont propres, comme la maîtrise du tour de parole (Kerbrat-Orecchioni, 1990).

Un robot doté de la parole est équipé d'une interface homme-machine qui le rend capable, dans une certaine mesure, de dialoguer avec un être humain. Autrement dit, il a la capacité de prendre la place d'un interlocuteur humain et, ce faisant, d'interagir avec un être humain en étant alternativement locuteur et allocutaire. De ce point de vue, l'enchaînement des messages entre un robot et un humain est plus ou moins long, à l'instar d'une conversation entre deux personnes. Pour permettre à un robot de communiquer, oralement ou par écrit, l'interface homme-machine doit maîtriser les mécanismes 1, 2 et 3 qui sont à l'œuvre pendant un échange d'informations. C'est le cas lorsque le robot se contente de satisfaire à une demande

d'information (Buvet *et al.*, 2019). Une conversation entre un robot et un être humain nécessite de maîtriser également le mécanisme 4. Ce mécanisme est constitué des règles qui gouvernent la succession des messages entrants et des messages sortants

Les quatre mécanismes langagiers sous-jacents à l'échange d'information entre un robot et un être humain sont généralement pris en charge par un système de dialogue constitué d'autant de modules³. Les messages entrants sont interprétés par un module qui analyse leurs contenus. Un module de correspondance associe les représentations des messages entrants à celles des messages sortants. Les messages sortants sont fournis par un module qui les produit. Un module de gestion des dialogues prend en charge la coordination du flux des messages. Il existe différents types de système de dialogue. Deux paramètres permettent de les distinguer: d'une part, le traitement des données et, d'autre part, leur architecture.

Plusieurs approches sont possibles pour le traitement des données. Les méthodes non dynamiques consistent à catégoriser les messages entrants par identification de mots clés puis à les associer avec des messages sortants préalablement enregistrés. Les systèmes de dialogue fondés sur de telles méthodes sont peu performants car les résultats qu'ils fournissent dépendent exclusivement d'informations pré-écrites. Certains de ces systèmes améliorent leur performance en exploitant des messages partiellement remplis que les mots clés viennent ensuite compléter (Gouritin, 2018).

Les méthodes semi-dynamiques reposent sur les principes de la recherche d'information. Des messages sortants sont préalablement stockés et indexés, puis associés aux messages entrants par le biais de

³ Nous négligeons ici la description des modules de speech-to-text et de synthèse vocale qui sont à l'entrée et à la sortie du système.

calculs de similarité (Bateman; Hovy, 1991). Les systèmes de dialogue basés sur ces méthodes sont plus performants que les précédents car l'analyse des messages entrants n'exige aucun prérequis. Ces systèmes sont adaptés pour les systèmes de question-réponse, ils le sont moins lorsqu'il s'agit de simuler des dialogues.

Les méthodes dynamiques sont de deux ordres. En premier lieu, il y a les méthodes qui se fondent exclusivement sur l'intelligence artificielle en exploitant des algorithmes de *deep learning*. A partir d'un panel de messages entrants associés de façon appropriée à des messages sortants, il s'agit d'entraîner un système de telle sorte qu'il produise automatiquement des messages sortants à partir de l'analyse des messages des messages entrants. Les systèmes qui s'appuient sur de telles méthodes obtiennent des résultats satisfaisants, notamment pour des domaines très contraints. Ils sont néanmoins tributaires de la qualité et de la quantité des données qui servent à entraîner le système (Cambrai, 2017).

En second lieu, il y a les méthodes qui font appel également à l'intelligence linguistique (Buvet, à paraître). Ces méthodes s'appuient sur des savoirs relatifs à la compréhension du langage, du point de vue de son mode de fonctionnement et de ses propriétés. Elles permettent de développer des outils informatiques qui exploitent les descriptions formelles issues de ces savoirs pour le traitement de l'information non structurée.

Les deux approches ont vocation à être complémentaires. L'intelligence linguistique apporte les connaissances qui servent à structurer les calculs effectués par des algorithmes. L'intelligence artificielle met à disposition sa puissance de calcul pour les dispositifs fondés sur l'intelligence linguistique. La méthode supervisée utilisée dans le cadre du machine/*deep learning* illustre ce point de vue.

Deux architectures modulaires des systèmes de dialogue sont possibles: une architecture circulaire et une architecture tubulaire. Dans la première architecture, la chaîne de traitement de l'information passe successivement par différents modules, un module servant de point d'entrée, un autre module de point de sortie. L'activation des modules s'effectue de manière linéaire en fonction du traitement de l'information. Dans la seconde architecture, les modules fonctionnent en parallèle. La chaîne de traitement de l'information est gérée par un module principal, celui qui gère les tâches en rapport avec le déroulement de la conversation. Les autres modules sont activés en fonction des besoins identifiés par le module principal. Des allers-retours s'effectuent entre ce module et les autres modules.

Le système de dialogue présenté ici repose sur une approche du type intelligence linguistique et sur une architecture modulaire du type tubulaire.

2. MODÉLISATION

La conception de l'échange d'information est de deux ordres: l'une est centrée sur son caractère élémentaire et l'autre sur son caractère global. Dans le premier cas de figure, il s'agit de l'émission et de la réception d'un message. Dans le second cas de figure, il est question de l'enchaînement des messages sous forme de dialogue. La notion d'énoncé est au cœur de la première conception, celle de conversation est centrale dans la seconde conception. Le point de vue conversationnel englobant le point de vue énonciatif, c'est ce dernier qu'il convient d'étudier dans un premier temps.

La sémantique est centrale dans les cinq mécanismes sous-jacents à l'échange d'information, *cf. supra*. De ce point de vue, est

fondamentale la notion de contenu propositionnel car elle contribue à expliquer l'encodage et le décodage d'un énoncé. Ce dernier a deux dimensions: une dimension sémantique, une dimension morphosyntaxique. La première dimension se manifeste dans un contenu propositionnel, la seconde correspond à la forme de l'énoncé.

Le concept de prédicat, emprunté à la logique, est généralement sollicité pour définir une proposition. Ce concept a deux acceptions distinctes selon qu'il relève de logique aristotélicienne ou bien de la logique non-aristotélicienne. La première acception définit le prédicat comme un attribut, une propriété. Il s'ensuit que la proposition se conçoit comme l'application d'un prédicat à un sujet comme l'indique cette représentation binaire: **Proposition => Sujet + Prédicat** (Ducrot; Todorov, 1972). La seconde acception assimile le prédicat à une fonction, au sens algébrique du terme. Il en résulte la représentation suivante de la proposition: **Proposition = Prédicat (Argument)** (Blanche, 1970). Les prédicats sont donc équivalents à des fonctions dont les variables sont les arguments. Autrement dit, les propositions correspondent à la projection des arguments sur les prédicats. Il est question ici de la deuxième acception du terme prédicat et de la conception de la proposition qu'elle implique, c'est-à-dire qu'elle dérive d'une structure prédicat-argument (Blanco; Buvet, 2009).

Un contenu propositionnel est défini comme une occurrence d'une structure prédicat-argument, telle que le prédicat correspond à une relation orientée entre des entités, qui sont assimilées à ses arguments. Par exemple, la structure prédicat-argument TRANSFERT_FINANCIER (ETRE_HUMAIN1 / ORGANISME/ CONTRIBUTION1, PRODUIT2, ETRE_HUMAIN3/ORGANISME3) est la représentation langagière d'une relation orientée, symbolisée par le prédicat TRANSFERT_FINANCIER, entre trois entités symbolisées

respectivement par ETRE_HUMAIN ou bien ORGANISME en ce qui concerne les arguments 1 et 3 et PRODUIT ou bien CONTRIBUTION pour ce qui est de l'argument 2. Les structures prédicat-argument constituent un savoir linguistique partagé par les membres d'une communauté langagière (Mejri, 2016). Les structures prédicat-argument sont de nature logico-sémantique. A la structure prédicat-argument TRANSFERT_FINANCIER (ETRE_HUMAIN1/ORGANISME1, PRODUIT/CONTRIBUTION2, ETRE_HUMAIN3/ORGANISME3), sont rattachées des énoncés comme *Monsieur Dupont achète une Peugeot à monsieur Durand, la société Martin vend une part de ses actifs à la société Dubois, Monsieur Lefevre paie sa taxe immobilière à l'État*. Ces différents énoncés partagent la même structure prédicat-argument mais diffèrent quant à leur contenu propositionnel. La situation de communication explique ces différences. Elle est déterminée par différents paramètres liés à l'énonciateur et aux conditions de son énonciation, notamment, le temps et le lieu de l'énonciation (Charadeau, 1983). La représentation fonctionnelle d'une structure-prédicat s'applique aux contenus propositionnels qu'elle subsume mais le prédicat est typé et les arguments sont instanciés: ACHAT (Monsieur Dupont1, Peugeot2, Monsieur Durant3); VENTE (société Martin1 , actif2, société Dubois3); PAIEMENT (Monsieur Lefevre1, taxe immobilière2, Etat3). Le prédicat TRANSFERT_FINANCIER englobe les types sémantiques, ACHAT, VENTE, PAIEMENT et ses arguments ETRE_HUMAIN1, ETRE_HUMAIN3, ORGANISME1, ORGANISME3, PRODUIT2 et CONTRIBUTION2 sont instanciés par les différentes dénominations des représentations. Un même contenu propositionnel peut donner lieu à des énoncés formellement différents: *Monsieur Dupont, achète une Peugeot à monsieur Durand; Monsieur Dupont, achète-il une Peugeot à Monsieur Durand?; Monsieur Dupont achèterait une Peugeot à*

monsieur Durand. Le caractère respectivement injonctif, interrogatif et hypothétique de ces énoncés les rend différents de l'énoncé purement informatif *Monsieur Dupont achète une Peugeot à monsieur Durand*. Ces nuances sémantiques sont la manifestation de l'intention communicative du locuteur.

Le décodage d'un message est une opération qui vise à interpréter le contenu propositionnel sous-jacent à un énoncé compte tenu de la situation de communication et de l'intention communicative du locuteur. L'analyse du décodage consiste ici à associer des énoncés à la représentation logico-sémantique de leurs contenus propositionnels. Elle ne se limite pas pour autant à traiter sa dimension explicite, elle prend en compte également sa dimension implicite. L'encodage d'un message est une opération qui a un contenu propositionnel, une situation de communication et l'intention communicative du locuteur comme point de départ et la forme d'un énoncé comme point d'arrivée. L'analyse de l'encodage consiste à expliciter les règles permettant d'aboutir à un énoncé bien formé à partir d'un contenu propositionnel. Quatre des cinq compétences langagières interviennent dans les processus de décodage et d'encodage: la compétence lexicale, la compétence grammaticale, la compétence sémantique, la compétence pragmatique.

Qu'il s'agisse du processus d'encodage ou du processus de décodage, l'énonciation peut s'inscrire dans un cadre plus large, celui des interactions verbales (Sacks, 1992). C'est notamment le cas des conversations. Il s'agit alors non seulement d'interpréter et de formuler des messages mais aussi de le faire d'une manière cohérente afin de favoriser l'échange d'informations. Un tour de parole s'instaure entre deux interlocuteurs avec la mise en place d'une relation dialogique qui se manifeste à travers le croisement de leurs propos respectifs. Les paramètres à prendre en compte pour l'analyse d'un

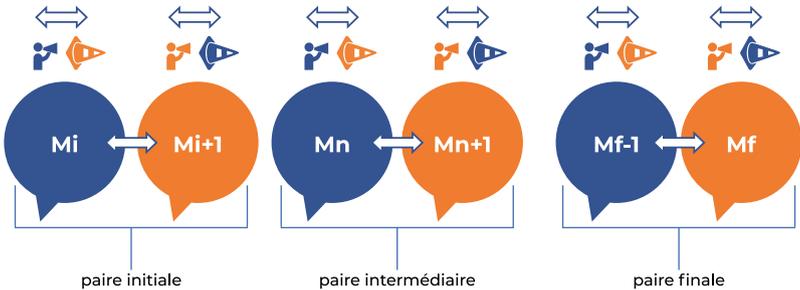
dialogue sont: 1) «le but conversationnel»; 2) «le déroulement»; 3) «le thème»; 4) «l'arrière-plan» (Caelen, 2003). Le paramètre 1 permet d'établir à quelle fin est mené le dialogue. L'objectif est nécessairement commun aux deux acteurs de la conversation, faute de quoi elle donne lieu à des malentendus et génère de l'incompréhension. Le paramètre 2 consiste à identifier les différentes étapes du dialogue, en particulier les étapes faisant l'objet de conventions sociales comme celles qui correspondent à la façon d'entamer une conversation et de la terminer. Le paramètre 3 sert à préciser dans quel cadre général s'inscrit le dialogue. Le paramètre 4 correspond au contexte. Le thème et l'arrière-plan ne sont pas de même nature. Le premier fait référence à un domaine de connaissance, le deuxième à la situation de communication. Le cadre général est obligatoirement partagé par les interlocuteurs, c'est moins le cas pour le contexte. Il est partagé lorsqu'il fait intervenir des éléments d'information extralinguistiques pour interpréter les énoncés. Il ne l'est pas obligatoirement lorsqu'il s'agit de préciser ce qui a occasionné la conversation. Son initiateur, en prenant la parole en premier, a potentiellement un objectif personnel, une intention dialogique, qui se révélera au fur et à mesure des échanges d'informations. Par conséquent, l'intention dialogique ne recouvre pas nécessairement le but conversationnel⁴.

L'unité minimale du dialogue est une paire des propos émis à tour de rôle par les interlocuteurs (Sachs, 1984). L'articulation entre les deux propos est fondée sur leur complémentarité, le contenu d'un propos s'inscrivant nécessairement dans le prolongement du propos qui le précède. Dans une unité dialogique, les interlocuteurs sont mobilisés au niveau de chaque propos. Il y a successivement encodage et décodage du premier propos puis encodage et décodage du second

⁴ La maïeutique de Socrate est une bonne illustration de cet aspect du dialogue, cf. *Le Théétète* de Platon.

propos. Le déroulement des propos progressent donc deux fois moins vite que celui des échanges d'information qui impliquent pour chaque propos, d'une part, du point de vue de locuteur, la formulation du message (le message sortant) e, d'autre part, du point de vue de l'allocutaire, l'interprétation du message (le message entrant). Le noyau de l'interaction est la continuité qui s'effectue entre le message sortant et le message entrant. Le schéma ci-dessus résume ces différents aspects du dialogue⁵.

Figure 1 – Les différents aspects du dialogue



Source: élaboration de l'auteur

L'enchaînement des paires de propos s'effectue dans un cadre coopératif, même dans des situations de confrontation ou de concurrence entre les participants. Une conversation repose sur l'engagement mutuel de deux interlocuteurs car «l'interaction (est un) processus complexe de coordination des actions (tel que...) en relation de co-présence, les participants à une interaction se rendent mutuellement intelligibles le sens de leurs actions et la

⁵ Remarquons que la représentation des paires ne recouvre pas toutes les situations possibles. Par exemple, un même locuteur peut enchaîner deux propos de suite avec une pause au milieu. Il s'ensuit qu'une paire peut être constituée d'un propos et d'une pause. Par ailleurs, le propos final n'est pas toujours un propos de l'interlocuteur qui a initié le dialogue.

compréhension qu'ils ont de ce qui se passe» (De Fornel; Leon, 2000). Sur la base de ce contrat dialogique, la conversation se développe graduellement, chaque propos répliquant au précédent. Une réplique est parfois imputable au même locuteur. Elle implique généralement une pause entre deux propos (Sacks; Schegloff, 1973). C'est notamment le cas lorsqu'elle relève de la fonction phatique de la communication (Jakobson, 1963). Il s'agit alors pour le locuteur de s'assurer auprès de l'allocutaire de la bonne transmission de l'information. Des propos plus ou moins conventionnels prennent en charge cet aspect de la conversation, par exemple *Vous m'avez compris?* Inversement, un allocutaire en se transformant en locuteur est parfois conduit à demander des clarifications sur le propos précédent (Ducan, 1972).

Le point de départ d'une réplique est la compréhension de la réplique précédente en tant que message entrant. Son point d'arrivée est sa formulation, en tant que message sortant. C'est donc à ce niveau qu'intervient le mécanisme de la mise en relation entre un message entrant et un message sortant. Ce dernier se doit d'être approprié au premier afin d'assurer le caractère interactif des propos qui s'enchaînent. Par exemple, un propos impliquant un choix à faire doit être complété par un propos faisant état de son engagement par rapport à l'une des alternatives proposées ou bien de son hésitation, voire de son refus de choisir. De ce point de vue, la compétence sémantique est fondamentale pour la mise en relation des messages. Les paires de propos et leur enchaînement sont les seuls observables dont dispose l'analyse conversationnelle (Sacks, 1984). Leur étude a fait apparaître des invariants, notamment lorsqu'il s'agit d'amorcer une discussion, de la conclure et de vérifier que la communication est établie et qu'elle continue de l'être durant toute la conversation. Par ailleurs, une conversation est régie par des principes de cohérence et

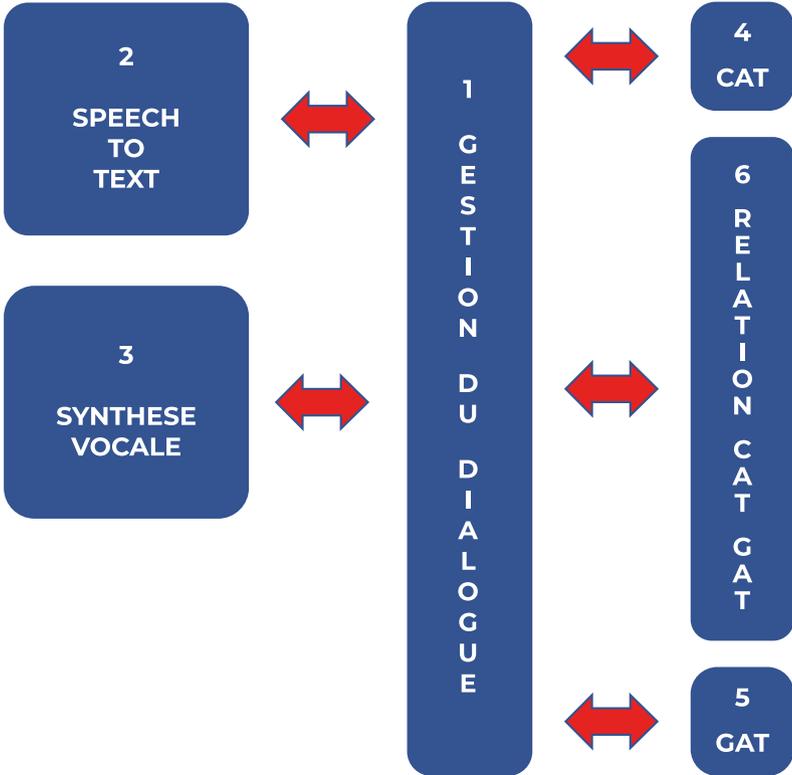
de progression thématique (Charaudeau, 1992). Enfin, elle s'appuie sur des stratégies discursives qui dépendent de l'intention communicative du locuteur qui a initié la conversation (Vanderveken, 1988). Ces paramètres sont traités au niveau du mécanisme de la coordination du flux conversationnel.

Les différents aspects de la modélisation sont pris en compte, à des degrés divers, dans le système de dialogue présenté dans la section suivante.

3. ASPECTS TECHNOLOGIQUES

La version 3 du système de dialogue est constituée de six modules: 1) un module de gestion de dialogues qui occupe une place centrale dans le système; 2) un module de speech-to-text; 3) un module de synthèse vocale; 4) un module de compréhension automatique de textes (CAT); 5) un module de génération automatique de textes (GAT); 6) un module de mise en relation de la sortie du module CAT et de l'entrée du module GAT. Le schéma ci-dessus représente l'architecture du système.

Figure 2 – Architecture du système



Source: élaboration de l'auteur

L'architecture est tubulaire. Le module de gestion du dialogue est le module principal. A ce titre, il gère entièrement la chaîne de traitement de l'information en mobilisant les autres modules en fonction des tâches à effectuer et compte tenu de leur spécialisation. L'information est traitée par le moyen d'allers-retours entre le module principal et les autres. L'information est traitée sur deux niveaux: le niveau élémentaire et le niveau global. Au premier niveau, ce sont les paires de propos qui sont analysées en tant qu'unités minimales du

dialogue. Au niveau global, c'est la succession des paires de propos qui est analysée.

L'ensemble du dialogue est donc séquencé en paires de propos dans le module 1. Deux cas de figure sont possibles selon que l'initiateur du dialogue est un être humain ou un robot. Dans le premier cas de figure, les principales tâches effectuées sont les suivantes: tâche a) le propos de l'utilisateur est capté dans le module 1 sous forme de fichier son et transféré au module 2 qui le transforme en fichier texte; tâche b) le fichier texte est envoyé au module 4 pour être associé à une représentation symbolique ; tâche c) cette représentation est communiquée au module 6 afin de lui associer une autre représentation symbolique; tâche d) la nouvelle représentation est communiquée au module 5 qui fournit un nouveau fichier texte ; tâche e) ce fichier texte est transmis au module 3 afin qu'il soit transformé en fichier son, activé par le module 1. Les cinq tâches permettent de compléter le propos initial par un nouveau propos, formulé par le robot. Le module 1 gère ensuite l'enchaînement des paires de propos selon les mêmes principes tout en veillant à la cohérence du dialogue jusqu'à son achèvement. Dans le second cas de figure, les tâches a), b) et c) sont supprimées pour le propos initial. Une représentation symbolique est envoyée directement au module 5 de telle sorte que les tâches e) et d) fournissent un propos initial formulé par le robot. Ensuite, à partir du propos suivant, formulé par l'utilisateur, les paires de propos et leur enchaînement sont gérés de la même façon par le module 1.

Le fonctionnement des modules CAT et GAT repose sur des ressources linguistiques exhaustives. Le système de compréhension automatique, incorporé dans le système de dialogue, exploite trois types de ressources linguistiques: 1) des dictionnaires électroniques; 2) des grammaires locales; 3) des ontologies. Les ressources du type

1) sont des listes de mots pleins associés à des métainformations linguistiques standardisées (Buvet; Grezka, 2009). Par exemple, le dictionnaire électronique des parties du corps recense l'ensemble des items lexicaux⁶ que subsume ce concept et associe chaque item à une description normalisée faisant état de sa catégorie grammaticale et de sa classe sémantique. Les ressources du type 2) sont des descriptions morphosyntaxiques très complètes de contenus propositionnels (Buvet, 2019). En tant que représentations formelles d'éléments cotextuels, les grammaires locales correspondent à la classe des grammaires les plus simples dans la hiérarchie de Chomsky (Chomsky; Miller, 1968). Elles se caractérisent par leur puissance combinatoire. Les grammaires locales font appel à des dictionnaires électroniques pour traiter les données lexicales. Elles sont implémentées sous formes d'automates qui parcourent les textes pour identifier des informations. Les ressources du type 3) servent à décrire explicitement un domaine de connaissance extralinguistique en indiquant le vocabulaire qui s'y rapporte (Guarino *et al.*, 2004). Les ontologies sont complémentaires des dictionnaires électroniques et des grammaires locales. Elles contribuent au traitement de l'implicite dans les discours⁷. Les trois types de ressources servent à reproduire les compétences lexicales, grammaticales, sémantiques et pragmatiques d'un être humain du point de vue du décodage de l'information. L'exploitation conjointe des trois types de ressources au niveau de ce module corrobore l'interdépendance de ces compétences langagières.

Le module CAT a pour fonction d'associer une représentation symbolique à un message entrant formulé par l'interlocuteur

⁶ Les items lexicaux sont des mots simples (par exemple, *main*) ou des mots composés (par exemple *avant-bras*), des mots courants (*tête*)

⁷ Les inférences sont calculées avec des règles qui exploitent des ontologies en fonction de la situation de communication. Par exemple, lorsque le locuteur humain dit *J'ai soif*, l'interlocuteur robot doit comprendre *Il faut me servir à boire*.

humain. La représentation du message s'appuie sur le formalisme descriptif des contenus propositionnels en termes de prédicat et d'argument, *cf. supra*. Elle prend en compte les caractères explicite et implicite des messages. Par exemple, le message entrant *J'ai soif* est représenté comme suit: SENSATION_SOIF (INTERLOCUTEUR_HUMAIN) & INJONCTION (INTERLOCUTEUR_HUMAIN, INTERLOCUTEUR_ROBOT, DON (INTERLOCUTEUR_ROBOT, INTERLOCUTEUR_ROBOT, BOISSON)). La première partie de la représentation renvoie au message explicite et la seconde partie mentionne ce qu'il a d'implicite.

Les ressources linguistiques exploitées dans le module GAT sont stockées dans une base de données. Les différentes propriétés d'un énoncé sont décrites formellement dans les tables de la base. La table dédiée aux prédicats fonctionne comme une table-mère du fait de leur rôle central dans la représentation des contenus propositionnels, *cf. supra*. Elle comporte trois types de champs. Premièrement, il y a les champs directement en rapport avec la nature des prédicats: forme et catégorie grammaticale. Deuxièmement, il y a des champs qui pointent vers les tables en rapport avec les autres propriétés des énoncés. Ces tables filles sont conçues ici comme des caractéristiques des prédicats: distributions; construction; actualisation⁸. Troisièmement, il y a les champs qui spécifient les représentations symboliques qui sont au départ du processus de génération des messages sortants et leurs spécificités pragmatiques. Ces champs permettent d'identifier les descriptions des prédicats qu'elles mobilisent.

Une partie des tables filles fonctionnent à leur tour comme des tables mères. C'est notamment le cas de la table construction qui pointe vers la table reconstruction. Une construction correspond à

⁸ L'actualisation concerne ici la spécification des temps verbaux et la détermination des substantifs compte tenu de la situation de communication.

la forme syntaxique d'un contenu propositionnel. Elle est tributaire de la forme du prédicat constitutif du contenu propositionnel. Par exemple, le prédicat verbal *manger* a comme construction standard X0 V (X1), c'est-à-dire un sujet suivi d'un verbe et d'un complément d'objet direct facultatif: *Tom mange un gâteau/ Tom mange*. D'autres formes syntaxiques sont possibles: *Tom, il mange un gâteau/ Tom, il mange/ Le gâteau, Tom le mange/ C'est Tom qui mange un gâteau/ C'est un gâteau que Tom mange/...* Ces variantes de la construction standard sont des reconstructions. Elles sont décrites dans la table éponyme⁹. La table distribution fait état du domaine d'arguments d'un prédicat, c'est-à-dire avec quels arguments il se combine pour constituer un contenu propositionnel. C'est une table fille et une table mère car elle pointe vers la table argument. Cette dernière décrit toutes les formes possibles d'un argument décrit sémantiquement dans la table distribution.

Enfin, il y a des tables qui servent à traiter d'autres aspects de la formulation d'un énoncé. C'est le cas de la table morphologie qui prend en charge l'analyse des mots variables (conjugaison des verbes, flexion des noms et des adjectifs) en fonction de leur cotexte. D'autres tables traitent des mots composés du point de vue de leur conjugaison ou de leur flexion ou prennent en charge les aspects ontologiques et pragmatiques de l'énonciation (par exemple les relations familiales en rapport avec un usager). Enfin, il y a des tables dédiées à des règles d'ajustement morphologique et d'ajustement syntaxique qui permettent en dernier ressort de réviser des formulations erronées du fait de la trop grande portée d'autres règles. Par exemple, une règle d'accord de l'adjectif avec le sujet explique l'énoncé erroné suivant: *La météo est au bell*. Cette erreur est corrigée avec la règle d'ajustement

⁹ Deux formes prédicatives qui partagent la même construction ne partagent pas toujours les mêmes reconstructions. Deux champs, l'un dans la table des prédicats, l'autre dans la table des reconstructions rendent compte de ces contraintes.

postérieure suivante: \$etre au belle-> \$etre au beau. L'application de cette règle aboutit à l'énoncé correct: *La météo est au beau*. Les relations entre les tables modélisent l'interdépendance entre les compétences lexicale, grammaticale, sémantique et pragmatique, cf. *supra*.

Le module de génération automatique de textes a pour fonction de produire un message sortant, sous forme de texte, à partir d'une représentation symbolique. La formulation de l'énoncé est effectuée par un algorithme qui s'appuie sur les descriptions des faits de langue enregistrées dans la base de données. Les principes qui ont guidé le développement de l'algorithme sont les suivants (Buvet *et al.*, 2019): 1) principe prédicatif: il s'agit d'établir si la représentation symbolique concerne la prédication simple ou la prédication complexe¹⁰; 2) principe distributionnel: les propriétés sémantiques du prédicat et de ses arguments sont stipulées; 3) principe pragmatique: les informations de nature ontologique et pragmatique sont exploitées pour encadrer le nombre de réponses possibles; 4) principe lexical: spécification de l'une des unités lexicales rattachées aux catégories sémantiques des prédicats et des arguments; 5) principe syntaxique: il s'agit de préciser la construction ou l'une des reconstructions des formes prédictives; 6) principe d'instanciation: les formes linguistiques, consécutives au principe 4, sont insérées dans les positions occupées par les prédicats et les arguments dans la construction ou la reconstruction, consécutive au principe 4; 7) principe d'actualisation: le temps prédicatif et la détermination argumentale sont précisés de telle sorte que les verbes soient conjugués, les noms et les adjectifs soient fléchis, les déterminants et les pronoms soient dotés d'un genre, d'un nombre et, le cas échéant, d'une personne. Ces principes s'appliquent de façon ordonnée.

¹⁰ La prédication complexe se distingue de la prédication simple en incluant une structure prédicat-argument dans son domaine d'arguments.

Ils sont complétés par des règles d'ajustement morphologique et syntaxique pour aboutir à des énoncés parfaitement formés. La très grande diversité des messages sortants permet au robot de simuler une conversation humaine. Cette diversité s'explique par la puissance combinatoire du système de génération qui s'appuie sur la variété lexico-syntaxique des prédicats et des arguments.

Le module de mise en relation de la sortie du module de compréhension et de l'entrée du module génération a pour fonction de traiter la cohérence des paires de propos en faisant en sorte que le message sortant soit cohérent avec le message entrant. Il gère l'interfaçage entre le décodage d'une information et l'encodage d'une autre information afin que la seconde s'inscrive soit dans le prolongement de la première en la complétant ou bien marque une rupture en respectant les conventions en usage. Il prend en charge également le traitement des inférences et de la paraphrase. Le rôle de ce module est fondamental au niveau élémentaire.

Le module de dialogue intervient au niveau global. Sa fonction est de gérer les conversations entre un robot et un être humain du début jusqu'à la fin. Il prend en compte aussi bien un dialogue initié par l'être humain qu'un dialogue initié par le robot en s'appuyant sur des scénarios conversationnels construits en fonction d'une thématique¹¹. La longueur des conversations est variable. Elle est limitée parfois à trois propos (typiquement une demande d'information émise par l'utilisateur suivi d'une réponse et d'un propos conclusif émis par la machine). Elle comporte généralement une dizaine de propos.

En l'état actuel des travaux, le robot sert à satisfaire deux types de demandes: des demandes d'information (par exemple *Que mange-*

¹¹ Par exemple, la thématique de la chambre donne lieu à des scénarios en rapport avec le nettoyage de la chambre, le déplacement vers la chambre, le chauffage de la chambre, la luminosité de la chambre, etc.

t-on ce soir?) et des demandes d'action (par exemple, *Allume la lumière*). Ces demandes et la satisfaction de ces demandes sont intégrées dans des scénarios qui comportent également des Items Exclusivement Conversationnels (IEC) formulés par le robot¹². Les IEC se répartissent dans deux catégories: d'une part, les items qui servent à encadrer une conversation; d'autre part, les items qui servent à s'assurer du bon déroulement de la conversation. La première catégorie regroupe des propos conventionnels qui contribuent à ouvrir une conversation (par exemple des salutations comme *bonjour* ou *comment allez-vous?*) ou bien à la clôturer (par exemple, *j'espère que ma réponse vous a satisfait* ou *je suis à votre disposition lorsque vous aurez besoin de moi*). La seconde catégorie concerne des propos qui invitent à mettre en place une conversation (par exemple, *que puis-je faire pour vous aider?*) ou bien à la poursuivre (par exemple, *souhaitez-vous parler d'autre chose?*), des propos correspondant à des formules de politesse (par exemple, *je vous remercie de l'avoir fait*), des propos qui vérifient la bonne transmission de l'information (par exemple, *est-ce que vous m'avez bien compris?* ou *suis-je assez clair?*) et des propos qui demandent des précisions. Ces derniers peuvent impliquer une reprise du propos précédent soit sous la forme d'un discours rapporté directement (*Vous avez dit: «je veux être raccompagné dans ma chambre»?*) soit sous la forme d'un discours rapporté indirectement (*Avez-vous demandé à être raccompagné dans votre chambre?*). Dans le second cas de figure, une reformulation, sous la forme d'une paraphrase, est également possible (par exemple *Voulez-vous que je vous emmène jusqu'à votre chambre?*). Les propos relatifs à une demande de répétition relèvent également de la seconde catégorie (par exemple *Pourriez-vous répéter ce que vous avez dit, s'il vous plaît, je n'ai pas bien compris?*).

¹² Un dialogue entre une machine et un être humain n'est pas strictement comparable avec un dialogue entre deux êtres humains.

Les scénarios implémentés dans le module de dialogue sont organisés en fonction des thématiques. Chaque thématique est traitée par plusieurs scénarios. Un scénario est constitué de messages entrants et de messages sortants. Les messages entrants résultent des propos émis par l'interlocuteur humain. Les messages sortants, sous leur forme audio, constituent autant de propos émis par l'interlocuteur robot en réaction à ceux de l'interlocuteur humain. Le traitement des messages entrants et des messages sortant s'effectue en faisant appel aux autres modules du système de dialogue, *cf. supra*. Il s'ensuit que les propos émis par un robot au sein d'un même scénario ont différents aspects du fait de la variété lexico-syntaxique inhérente aux énoncés produits par le module de génération. La puissance combinatoire du système de dialogue donne lieu à environ 30 000 formes de conversations différentes pour une même thématique¹³. La forme d'un scénario diffère selon que l'initiateur de la conversation est un être ou bien un être humain.

Il y a une asymétrie entre les propos tenus par le robot et ceux tenus par l'être humain car ce dernier n'est pas tenu par les conventions sociales qui caractérisent une conversation entre deux êtres humains. Typiquement, les propos en rapport avec l'ouverture ou la clôture d'une conversation sont souvent omis dans la composante conversationnelle imputable à l'interlocuteur humain. De même, d'une façon générale, il ne recourt pas aux formules de politesse. De tels propos sont nécessairement tenus par le robot puisque sa conversation simule celle d'un être humain et, de ce fait,

¹³ La formule pour obtenir ce résultat est celle d'un arbre complet: $MFC = ((MST)^{\max(MPS/2, MVP) - 1})$ avec MFC signifiant Moyenne des Formes Conversationnelles, MST signifiant Moyenne des Scénarios dans un Thème, MPS signifiant Moyenne des Propos dans un Scénario, MVP signifiant Moyenne des Variantes dans un Propos. En l'occurrence MST = 8, MPS = 5 et MVP = 6 donc MFC = 32767. Le fait que seuls les propos du robot sont calculés explique la division par 2 de MPS (1 propos est imputable au robot, l'autre à son vis-à-vis humain).

s'appuie sur les conventions qui régissent le déroulement d'une conversation. Toutefois, il n'est pas exclu qu'un être humain s'adresse à un robot humanoïde comme il s'adresserait à un être humain dans la mesure où, conçu pour être un compagnon de vie, le robot fait l'objet d'une humanisation (Tisseron, 2015).

4. PERSPECTIVES

La question des ressources linguistiques est centrale dans le système de dialogue présenté ici. Son fonctionnement est tributaire de la qualité des descriptions des faits de langue, notamment dans les modules CAT et GAT. C'est d'autant plus vrai que les ressources ne sont pas de même nature dans les deux modules. Le fonctionnement du module CAT s'appuie un moteur d'analyse sémantique performant qui exploite des descriptions exhaustives du français, entre autres langues, sous la forme de grammaires locales et de dictionnaires électroniques. Le fonctionnement du module GAT repose sur un algorithme qui exploite des descriptions stockées sous la forme d'une base de données. Ce sont principalement ces dernières descriptions qui doivent être étendus.

Les langues ont la particularité à partir d'un nombre fini d'unités linguistiques de produire un nombre infini d'énoncés (Mejri, 2019). La puissance combinatoire explique cette productivité discursive (Danon-Boileau, 2007). Celle-ci s'explique en partie par l'interdépendance entre les principales compétences langagières, notamment la compétence sémantique, la compétence lexicale et la compétence grammaticale. Cette interdépendance est modélisée dans la base de données de telle sorte qu'est décrit le comportement syntactico-sémantique de chaque unité lexicale fonctionnant comme un prédicat. Cette description est utilisée dans le module CAT.

Pour rendre le système de dialogue plus performant, il doit traiter rapidement de nouvelles thématiques conversationnelles. De ce point de vue, il est impératif que la description des prédicats soit améliorée au niveau de la base de données. Aussi, nous sommes en train de développer, dans le cadre de l'intelligence étendue, un système d'apprentissage qui s'appuiera sur les descriptions existantes pour identifier automatiquement de nouveaux prédicats ainsi que leur comportement syntactico-sémantique. Ce faisant, la base de données sera enrichie de nouvelles descriptions.

RÉFÉRENCES

AUSTIN, J. **Quand dire c'est Faire**. Paris: Seuil, 1962.

BATEMAN, J.; HOVY, E. Computers and Text generation: principles and uses. *In*: BUTLER, C. (Ed.) **Computers and texts: an applied perspective**. Oxford: Blackwell, 1991.

BLANCHE, R. **La logique et son histoire d'Aristote à Russell**. Malakoff: Armand Colin, 1970.

BLANCO, X.; BUVET, P. A. Présentation: Les représentations des structures prédicat-arguments. **Langage**, Paris, v. 176, p. 3-6, 2009.

BLANCO, X. Valeurs grammaticales et structures prédicat-argument. **Langages**, Paris, v. 176, p. 31-42, 2009.

BRACOPS, M. **Introduction à la pragmatique**. Paris: De Boeck Supérieur, 2010.

BUVET, P. A.; FACHE, B.; ROUAM, A. How Does a Robot Speak? About the Man-Machine Verbal Interaction. *In*: INTERNATIONAL WORKSHOP APPLICATIONS OF KNOWLEDGE REPRESENTATION AND SEMANTIC TECHNOLOGIES IN ROBOTICS, 3., 2019, Macau. **Proceedings** [...]. Suwon: AnSWeR, 2019. p. 79-93. Disponível em: https://swb.skku.edu/_res/cnrlab/etc/Joint_Proceedings_of_SDMM19_and_AnSWeR19.pdf.

BUVET, P. A.; GREZKA, A. Les dictionnaires électroniques du modèle des classes d'objets. **Langages**, Paris, v. 176, p. 63-79, 2009.

- BUVET, P. A. Comment parle un robot? **Linguística e Semiótica**, v. 23, n. 45, p. 30-58, 2019.
- BUVET, P. A. (no prelo). Linguistique et intelligence. **Linguistique et...**, Peter Lang.
- CAELEN, J. Stratégies de dialogue. *In*: HERZIG, A.; CHAIB-DRAA, B.; MATHIEU, P. (Eds.). **MF03: modèles formels de l'interaction - Actes des Secondes Journées Francophones**. Lille: Cépaduès, 2003.
- CAMBRAI, T. **L'intelligence artificielle expliquée: comment les algorithmes et le Deep Learning dominant le monde**. Paris: Albin Michel, 2017.
- CHARAUDEAU, P. **Langage et discours: éléments de sémiolinguistique**. Vanves: Hachette, 1983.
- CHARAUDEAU, P. **Grammaire du sens et de l'expression**. Vanves: Hachette, 1992.
- CHOMSKY, N.; MILLER, A. **L'analyse formelle des langues naturelles**. Paris: Gauthier-Villars, 1968.
- DANON-BOILEAU, L. **Le sujet de l'énonciation**. Paris: Orphys, 2007.
- DE FORNEL, M.; LEON, J. L'analyse de conversation, de l'ethnométhodologie à la linguistique interactionnelle. **Histoire Épistémologie Langage**, v. 22, n. 1, p. 131-155, 2000.
- DUCAN, S. Some signals and rules for taking speaking turns in conversations. **Journal of Personality and Social Psychology**, v. 23, n. 2, p. 283-292, 1972.
- DUCROT, O.; TODOROV, P. **Dictionnaire encyclopédique des sciences du langage**. Paris: Seuil, 1972.
- GOURITIN, T. L'arnaque chatbots durera-t-elle encore longtemps? 2018. Disponível em: <https://www.frenchweb.fr/larnaque-chatbots-durera-t-elle-encore-longtemps/305697>. Acesso em: 24 ago. 2023.
- GROSS, M. Les bases empiriques de la notion de prédicat sémantique. **Langages**, Paris, v. 63, p. 7-52, 1981.
- GUARINO, N.; OBERLE, D.; STAAB, S.; VOLZ, R. Developing and managing software components in an ontology-based application server. *In*: JACOBSEN, H.-A. (Ed.) **Middleware 2004 - ACM/IFIP/USENIX - International Middleware Conference Toronto, Canada, October 18-22, 2004**. Proceedings. Berlin: Springer, 2004.

- HARRIS, Z. S. **Notes du cours de syntaxe**. Paris: Le Seuil, 1976.
- JAKOBSON, R. **Essais de linguistique générale**. Paris: Les éditions de minuit, 1963.
- KERBRAT-ORECCCHIONI, C. **Les interactions verbales**. t. 1. Malakoff: Armand Colin, 1990.
- MARTIN, R. **Linguistique de l'universel**. Réflexions sur les universaux du langage, les concepts universels, la notion de langue universelle. Paris: Académie des Inscriptions et Belles-Lettres, 2016.
- MEJRI, S. Le prédicat et les trois fonctions primaires. *In*: COSTA, D. de S. S.; BENÇAL, D. R. **Nos caminhos do léxico**. Campo Grande: Editora UFMS, 2016. p. 313-337.
- MEJRI, S. De l'inarticulé dans le langage. **Les cahiers du dictionnaire**, v. 11, p. 25-58, 2019.
- MUFWENE, S. L'émergence de la complexité langagière du point de vue de l'évolution du langage. **La Clé des Langues**, ENS de LYON/DGESCO, 2010. Disponible em: <http://cle.ens-lyon.fr/plurilingues/langue/le-changement-linguistique/l-emergence-de-la-complexite-langagiere-du-point-de-vue-de-l-evolution-du-langage>. Acesso em: 24 ago. 2023.
- PARISSE, C. La morphosyntaxe: Qu'est-ce qu'est ? - Application au cas de la langue française. **Rééducation orthophonique**, v. 47, n. 238, p. 7-20, 2009.
- PIERREL, J. M. **Ingénierie des langues**. Paris: Lavoisier, 2000.
- SACKS, H.; SCHEGLOFF, E. A.; JEFFERSON, G. A simplest systematics of the organization of turntaking in conversation. **Language**, Paris, v. 50, n. 4, p. 696-735, 1974.
- SACKS, H. **Lectures on conversation**. Oxford: Blackwell, 1992.
- SACKS, H. Perspectives de recherche. **Problèmes d'épistémologie en sciences sociales**, v. 3, p. 138-44, 1984.
- SILBERZTEIN, M. **Formalizing Natural Languages**. London: ISTE-Wiley, 2016.
- TISSERON, S. **Le jour où mon robot m'aimera**. Paris: Albin Michel, 2015.
- VANDERVEKEN, D. **Les Actes de discours**. Bruxelles: Mardaga, 1988.

ANALOGIA E DESAUTOMATIZAÇÃO: PROCESSOS NEOLÓGICOS CONVERGENTES

Aderlande Pereira Ferraz

INTRODUÇÃO

Quando se percebe no léxico uma nova forma, uma nova acepção associada a uma forma já existente ou um estrangeirismo importado de outro sistema linguístico, tem-se, para cada caso, um neologismo figurando como uma palavra nova. Como toda língua em evolução apresenta variações e mudanças em todos os seus constituintes, a obsolescência de unidades lexicais que caem em desuso e o surgimento de novas unidades no léxico, os neologismos, são ocorrências naturais, previstas nos padrões de estruturação lexical das línguas vivas, como salienta Ferraz (2008).

Ao tratar do neologismo necessário em face do neologismo expressivo, Guilbert (1975, p. 40-41) os distinguiu como “neologismo denominativo” e “neologismo estilístico”, respectivamente. O primeiro diz respeito à necessidade de exprimir conceitos ou nomear realidades novas (ideias, objetos, produtos, serviços); o outro é uma criação neológica estilística, motivada pela busca de maior expressividade do discurso.

Partindo dessas considerações, nosso objetivo é apresentar alguns resultados da pesquisa desenvolvida no âmbito da Faculdade de Letras da Universidade Federal de Minas Gerais, da qual se originou o *Observatório de neologia em textos publicitários*¹. Com tal *Observatório*, textos publicitários veiculados pelas revistas noticiosas *Veja*, *Isto é* e *Época*, bem como em circulação na mídia eletrônica (sítios da internet e redes sociais), têm sido examinados criteriosamente desde 2006, com o propósito de identificar, recolher e analisar os itens lexicais candidatos a neologismos e, posteriormente, classificar as unidades consideradas neológicas, em consonância com as sucessivas edições das obras lexicográficas que compõem o *corpus* de exclusão estabelecido. Com isso, a publicidade de anúncios, seja por meio impresso, seja eletrônico, tem propiciado farto material de análise, caracterizando o *Observatório* como verdadeiro manancial de produtos neológicos, no que concerne à unidade básica do léxico.

Assim, o escopo do *Observatório* consiste em observar e analisar o fenômeno do surgimento de palavras novas em uso no português contemporâneo do Brasil, especificamente em textos publicitários, procurando difundir aspectos da inovação lexical que possam contribuir para o desenvolvimento da competência lexical, com abordagens pedagógicas sobre a neologia.

O estudo da competência lexical, a partir da coleta de neologismos, permite-nos saber como o usuário da língua forma unidades léxicas novas consideradas boas ou aceitáveis e evita a formação de unidades inaceitáveis, em consideração às restrições e bloqueios que ele encontra (Sandmann, 1991).

¹ OBSERVATÓRIO de neologismos na mídia publicitária eletrônica: banco de dados preparado por Aderlande Pereira Ferraz. Faculdade de Letras da Universidade Federal de Minas Gerais. Belo Horizonte, 2023. Acesso restrito.

Em face dos diversos processos neológicos em estudo no âmbito de nosso *Observatório*, destacamos os processos de analogia e desautomatização, de que resultam formações neológicas que se enquadram nas duas classes distinguidas por Guilbert (1975), já citadas, importando ressaltar que entre os neologismos estilísticos sob análise, alguns são casos de *hapax legomenon*, inteiramente efêmeros.

Não raro, a efemeridade impede que certas criações lexicais encontrem acolhida no sistema da língua, principalmente quando os próprios falantes não se sentem motivados à utilização ou à reutilização dessas criações. É que a frequência de uso dos neologismos faz com que, gradativamente, a sensação de novidade lexical vá se perdendo até que, naturalmente, as unidades neológicas passam a integrar o conjunto das unidades memorizadas e de distribuição regular entre os usuários da língua, o que favorece, em muitos casos, que essas unidades lexicais sejam registradas em dicionários de língua, caracterizando, assim, também a sua desneologização.

Como marco teórico, partimos do estudo seminal de Saussure sobre analogia, contido em seu *Curso de Linguística Geral* (1978 [1916]), retomando a noção de quarta proporcional para descrever certos padrões de formação de palavras.

Contudo, uma questão logo se impõe: no que concerne aos processos neológicos de que resulta a formação de palavras novas, que relação pode-se observar entre os processos de analogia e desautomatização?

Neste capítulo, a analogia e a desautomatização entram na pauta de estudo como processos neológicos convergentes.

A metodologia empregada para a realização deste trabalho apresentou os seguintes passos: seleção do *corpus* de análise (textos publicitários); estabelecimento dos critérios de detecção dos neologis-

mos; fixação de um *corpus* de exclusão como critério para identificar a neologicidade de uma unidade do léxico; extração e análise dos neologismos detectados; organização das informações, especialmente sobre os casos de analogia e desautomatização.

O repertório lexicográfico, composto especialmente por dicionários de língua, a constituir um dos critérios para a delimitação da unidade léxica neológica, figura aqui como um *corpus* de exclusão, como está explicitado em Ferraz (2020, p. 168):

Os dicionários escolares que constituem o acervo aprovado pelo Programa Nacional do Livro Didático (PNLD-Dicionários de 2012), classificados de tipo 4, integram atualmente o *corpus* de exclusão adotado pelo *Observatório*, por levar em conta o público a que se destinam. São quatro dicionários destinados aos estudantes do ensino médio: o *Dicionário Houaiss Conciso*, da editora Moderna; o *Novíssimo Aulete Dicionário Contemporâneo da Língua Portuguesa*, da editora Lexicon; o *Dicionário UNESP do Português Contemporâneo*, da editora Piá; e o *Dicionário da Língua Portuguesa Evanildo Bechara*, da editora Nova Fronteira. No âmbito do PNLD-Dicionários 2012, tais dicionários foram distribuídos às escolas públicas de todo o território brasileiro pelo Ministério da Educação, que reconheceu a importância do dicionário como instrumento didático nas aulas de língua portuguesa.

Nesse contexto, o dicionário de língua, pelo prestígio cultural de que goza e por ser um dos pilares do saber metalinguístico, é o filtro por meio do qual é formalmente atestada uma unidade lexical neológica em uso, assim como também é o marco de sua lexicalização, isto é, o final de um processo neológico, o que vem caracterizar, sob o critério lexicográfico, a sua desneologização.

1. A INOVAÇÃO LEXICAL NO GÊNERO PUBLICITÁRIO

Com o gênero publicitário, assistimos a uma grande manifestação de criação lexical no português do Brasil. Certamente, pelo caráter de novidade que cada anúncio deve trazer, o discurso publicitário se reveste da necessidade de utilização de palavras novas que, em grande parte, são criadas no interior do próprio discurso ou são adotadas quando revelam a contemporaneidade de uso pela comunidade linguística.

É nesse contexto que passamos a observar a produção de textos, especialmente os anúncios, em manifestação no gênero publicitário.

De acordo com o que salienta Ferraz (2010, p. 253),

Uma das razões, talvez a principal, que justificam a publicidade está na condição de o regime de produção de uma dada sociedade apresentar-se suficientemente desenvolvido para satisfazer mais que as meras necessidades materiais da sua população. E essa população, ou parte dela, para corresponder a tal expectativa, deve viver acima do nível da subsistência, isto é, deve estar em condições de consumir bens e serviços nem sempre necessários.

Neste trabalho, o gênero publicitário é compreendido como um discurso no qual o enunciador emite uma mensagem (por meio de anúncio publicitário) em direção ao enunciatário coletivo (potenciais consumidores), com o objetivo de induzir à aquisição de um produto ou serviço. Importa considerar que tal enunciatário não é um mero destinatário da comunicação, mas também produtor do discurso (cada potencial consumidor), pelo fato de ser a atividade de leitura um ato de linguagem que envolve uma dimensão dialógica. Isto é, o leitor

dialoga com o texto e com o contexto, assumindo um papel de sujeito, que procura e produz sentidos (Bakhtin, 2003, p. 175).

Pela realização de diferentes trabalhos sobre o gênero publicitário (Ferraz, 2010; 2019; 2020), temos observado que tal gênero age como uma grande força a incidir sobre a língua portuguesa, no sentido de forçar uma contínua transformação linguística, com reflexos em vários aspectos da língua e, com especial destaque, no campo lexical, com a criação de neologismos. Com argumentos sempre persuasivos, o discurso publicitário procura dar ênfase à qualidade supostamente superior de um produto ou serviço, utilizando, para isso, de todas as possibilidades expressivas que a língua oferece, principalmente a criação de neologismos. Estes se destacam por transmitir um significado específico e por chamar a atenção do receptor pela sensação de novidade que provoca.

É que a linguagem publicitária, fortemente caracterizada pela dinamicidade do léxico, sente-se permanentemente impulsionada a renovar-se, não apenas para nomear mercadorias, mas, sobretudo por tecer enunciados cheios de apelos estilísticos que procuram divertir, motivar, seduzir, fazer sonhar, entusiasmar, tornando, assim, a produção de neologismos um fator inevitável.

Importa ressaltar, entretanto, que é preciso de início definir um critério de reconhecimento do neologismo. No que diz respeito à gênese neológica e às circunstâncias em que esta ocorre, claro está que não é pelo fato de uma unidade léxica ter caráter inédito que passa a ser imediatamente considerada neológica.

Há três fases da neologia que devem ser observadas: (i) a fase inicial do processo, quando o neologismo está sendo criado; (ii) a fase que sucede à criação e se refere à recepção ou ao julgamento de sua

aceitabilidade por parte dos destinatários; (iii) a fase em que começa o processo de desneologização.

É, portanto, na terceira fase da neologia que a unidade léxica nova se encontra em condições de ser dicionarizada. Isso porque a comunidade linguística exerce o papel fundamental de determinar, pelo uso, a sua inserção no acervo lexical da língua.

O estatuto de neologismo é conferido aqui a uma unidade do léxico colhida de textos publicitários da mídia eletrônica quando esta passou do momento de criação ao momento de recepção e aceitabilidade pelos destinatários, sem ter sido ainda dicionarizada. Com isso, dois critérios neológicos estão sendo considerados, o psicolinguístico e o lexicográfico. O primeiro refere-se ao que Sablayrolles (2003) denomina ‘sentimento de neologicidade’, ou seja, trata-se da sensação que o usuário da língua demonstra diante de uma nova unidade; o segundo reporta-se à ausência de registro da unidade em repertórios lexicográficos, geralmente dicionários de língua geral. Nesse trânsito, a unidade léxica criada poderá ter seu emprego por vários locutores e será apresentada como um neologismo, identificado segundo o critério psicolinguístico, bem como pelo lexicográfico, fundamentado pela consulta a um *corpus* lexicográfico de exclusão, composto de obras de referência básica para o português brasileiro, já destacadas na seção anterior.

O dicionário de língua, pelo prestígio cultural de que goza e por ser um dos pilares do saber metalinguístico, é o filtro com maior grau de objetividade, por meio do qual é formalmente atestada uma unidade lexical neológica em largo uso, estabelecendo assim a sua lexicalização, isto é, o final de um processo neológico. Assim, para os neologismos recolhidos da linguagem publicitária eletrônica, tal critério será formalmente destacado pelo caráter de objetividade na obtenção dos dados, como especificado em Ferraz (2008, p. 154).

Partindo então da coleta e descrição de neologismos, veiculados por peças publicitárias da mídia eletrônica, em circulação em variados sítios e redes sociais na internet, não se pode deixar de reconhecer a celeridade com que se renovam os textos por ela veiculados, o que torna a publicidade digital muito dinâmica e produtiva, favorecendo o surgimento de novas unidades no léxico, tanto no que diz respeito às lexias simples quanto às compostas e complexas (Pottier, 1972, p. 67).

2. A NEOLOGIA FORMAL EM ESTUDO

A expansão lexical de uma língua geralmente ocorre por meio da reciclagem do material lexical já existente, isto é, a formação de palavras novas resulta, majoritariamente, do aproveitamento de fragmentos de material lexical, reconhecido por padrões gerais de estruturação, em novas construções. Com isso, o léxico, para incorporar unidades novas, recorre à neologia, que compreende três mecanismos, todos muito comuns: a neologia formal (construção de palavras por meio de regras do próprio sistema linguístico, com a utilização de procedimentos formais internos no nível morfológico, sintático e fonológico); a neologia semântica (expansão de sentido, quando da reutilização, com novos significados, de formas léxicas já existentes); e a neologia de empréstimos (importação de unidades léxicas de outros sistemas linguísticos, as quais se podem apresentar adaptadas ou não à nova língua).

Dos três mecanismos, interessa-nos a neologia formal por focalizar, especificamente, os processos de analogia e desautomatização.

A neologia formal tem sido muito produtiva no português pelos mecanismos morfemáticos, os quais permitem formar novas palavras a partir da combinação de morfemas, sejam estes bases ou

afixos. De acordo com o modo de ocorrência dos morfemas, diversos processos de formação de palavras são possíveis, quais os processos de prefixação, sufixação, composição, reduplicação etc. Nessa linha de desenvolvimento, a neologia formal abrange criações que decorrem da alteração de uma forma preexistente, por meio do espelhamento lexical, como nos casos de analogia e desautomatização.

Esses dois processos neológicos, como já salientados linhas atrás, embora raramente tratados pelas gramáticas tradicionais, são produtivos no português brasileiro contemporâneo, como se verificam em textos publicitários e jornalísticos da mídia eletrônica.

Como demonstração da produtividade desses processos tanto no gênero publicitário quanto no jornalístico, identificamos, como produtos neológicos resultantes do processo de analogia, os exemplos em (1) e (2), colhidos de manchetes jornalísticas da mídia eletrônica; assim como pelo processo de desautomatização, consideramos as formações neológicas em (3) e (4), oriundas de textos publicitários, veiculados em revistas noticiosas.

(1) QuarenPena, comunicação entre as pessoas neste tempo de pandemia.

(2) Carentena: Como sobreviver e sair bem dessa quarentena da Covid-19.

(3) O novo Credicard Cash Back é a novidade que vai dar o que gastar (Isto é, 13 ago. 2003, p. 17)

(4) Seu filho se diverte ao mesmo tempo que ganha uma mãozinha nos estudos (Veja, 22 mar. 2000).

Como se observa nos contextos de uso, especialmente nas unidades grifadas em (1) e (2), *quarenPena* e *carentena* são neologismos formados por analogia e remetem a quarentena. A primeira unidade léxica diz respeito ao blog de alguns dos alunos do Agrupamento de

Escolas de Ribeira de Pena (Portugal), cujo objetivo é apresentar, durante a quarentena, testemunhos, notícias e recomendações escritas pelos alunos da turma do 11º B, provocando interações entre os leitores do blog. A segunda é uma crítica ao custo alto que recairá sobre a sociedade, em face do impacto emocional na vida das pessoas que, pela falta de interação social (presencial), decorrente da quarentena imposta pela disseminação do novo coronavírus, podem sentir sofrimento mental: depressão, solidão etc.

Em (3), *dar o que gastar*, que faz um trocadilho com a expressão idiomática original “dar o que falar”, aparece em uma propaganda de cartão de crédito na qual o verbo ‘gastar’ ajusta-se inteiramente ao produto anunciado.

Em (4), *ganhar uma mãozinha*, o que se vê é a desautomatização fraseológica com a utilização de componentes que entram em relação de antonímia com o significado global da expressão idiomática “dar uma mãozinha”. Trata-se de ganhar/dar uma mãozinha (sentidos opostos), como se pode comparar com a expressão original em (5).

(5) O futuro a Deus pertence, mas você pode dar uma mãozinha. Para saber mais sobre o nosso trabalho, ligue 3679-2000 ou acesse www.goldeletra.org.br (Isto é, 10 jan. 2001, p. 83).

3. O PROCESSO DE ANALOGIA NO GÊNERO PUBLICITÁRIO

A analogia em estudo remonta à Antiguidade Clássica, em especial aos gramáticos gregos do século II a. C., os quais discutiam sobre a relevância do caráter de regularidade atribuído aos fenômenos linguísticos, presente no princípio da analogia.

Mais tarde, Saussure (1978, p. 271), ao tratar das relações lexicais estabelecidas pela palavra, põe em relevo a analogia, considerando-a como o princípio de criação na língua. O realce marcante às relações associativas e à analogia como processo que desencadeia a formação de palavras está muito evidente em sua obra, como se pode ver pelo seguinte trecho.

Poder-se-iam classificar as palavras de acordo com sua capacidade relativa de gerar outras conforme são elas próprias mais ou menos decomponíveis. As palavras simples são, por definição, improdutivas (cf. *armazém, árvore, raiz* etc.). *Armazenista* não foi gerado por *armazém*; foi formado sobre o modelo de *cambista – câmbio* etc. Do mesmo modo, *armazenar* deve sua existência à analogia com *forrar, limpar, pensar* etc., que contêm *forro, limpo, peso* etc (Saussure, 1978, p. 273).

O princípio da analogia está presente, ainda que de modo inconsciente para o usuário da língua, na criatividade como na produtividade (Bauer, 1983, p. 97). No que diz respeito às formações lexicais, a analogia deve ser identificada na relação que une as formas entre si, como salientou Saussure (1978), por meio da noção de quarta proporcional, em que se considera que $A - B = C - X$. Isto é, se os elementos A, B, e C são conhecidos, a dedução de X é automática, como se vê na exemplificação apresentada pelo grande linguista.

[...] para explicar a aparição de *honor* frente a *honôs*, temos que fazer apelo a outras formas, como o mostra a fórmula da quarta proporcional:

ōrātōrem — ōrātor = honōrem — X

X = honor

e esta combinação não teria qualquer razão de ser se o espírito não associasse pelo sentido as formas que a compõem (Saussure, 1978, p. 271-272).

A analogia está diretamente associada à criação por meio do espelhamento de uma forma anterior, entendendo-se que a formação analógica gerada não implica necessariamente o desaparecimento da forma que lhe serviu de fonte. Sabemos que *honos* e *honor* coexistiram por algum tempo até que a forma menos regular saiu de circulação, sobressaindo então *honor*. Desse modo, a analogia é um processo que garante a renovação lexical no âmbito da língua por meio da manutenção de elementos os quais reaparecem em combinações novas, o que induz à perenidade potencial das formas da língua, que se refazem analogicamente (Basílio, 1997, p. 16).

Não são raros no português os processos que operam sobre palavras pré-existentes para formar palavras novas, seja por meio da reanálise (desmembramento das partes) de formas já produzidas ou por uma operação sobre uma base específica, a fim de se produzir um produto de certo modo previsível, como em *brahmeiro*, a partir da regra de adição do sufixo *-eiro* a substantivos, como ilustra a peça publicitária a seguir.

Figura 1 – Anúncio publicitário



Fonte: Disponível em: <https://www.facebook.com/Os-Brahmeiros-759825497405972/>. Acesso em: 16 jun. 2021.

Desse modo, queremos ressaltar a atuação de um princípio de formação subjacente a um par de itens lexicalmente relacionados, a gerar formações analógicas, comumente espelhadas em compostos ou derivados, que não podem ser contabilizadas por qualquer tipo de regra, como bem salienta Bauer (1983, p. 96):

These are genuine analogical formations which cannot be accounted for by any kind of rule. It should be noted that this definition of an analogical formation is much narrower than the one usually implied in diachronic studies and is, in a sense, as closely related to creativity as to productivity.

Nesse sentido, o neologismo que aparece na peça publicitária a seguir pode ser explicado pelo processo clássico de analogia.

Figura 2 – Anúncio publicitário



Fonte: Disponível em: <https://www.facebook.com/McDonaldsBrasil/photos/a.304652009573356/2613390915366109/?type=3&theater>. Acesso em: 21 abr. 2020.

A formação analógica de *smartfome*, que diz respeito aos produtos do anunciante McDonalds associados à “fome inteligente” e remete à unidade já existente *smartfone*, pode ser observada na forma

que corresponde à incógnita da quarta proporcional (Saussure, 1978, p. 271-272), como em *Smart + (tele)fone: smartfone : smart + fome: x*, donde $x = \text{smartfome}$.

Essa mesma operação está presente na criação do neologismo *Mcshake*, que aparece no texto publicitário seguinte, ainda do mesmo anunciante, em cuja formação está o mecanismo da analogia. Trata-se da formação *Mcshake*, espelhada em *milk-shake*.

Figura 3 – Anúncio publicitário



Fonte: Disponível em: <https://www.facebook.com/McDonaldsBrasil/photos/a.358327450872478/2890197714352093/?type=3&theater>. Acesso em: 20 abr. 2020.

O item léxico *milk-shake* é um empréstimo lexical, registrado nos principais dicionários de língua portuguesa, figurando como uma unidade incorporada ao léxico do português brasileiro. Aproveitando esse contexto de uso, surge o neologismo *Mcshake*, que remete a *milk-shake*, nomeando um produto do anunciante McDonalds, em cuja gênese subjaz o princípio da analogia, de modo que a partir de *milk-shake* (bebida à base de leite batida com sorvete), *Mcshake* tem fácil interpretação.

Morfologicamente, para a formação de *Mcshake* ocorreu a junção de *McDonalds + milk-shake*, de modo que o produto gerado é resultante da segmentação aplicada a cada uma das formas pré-existentes. A

fusão de duas unidades do léxico em um todo fonético, apresentando apenas um único acento e aproveitando as propriedades semânticas das palavras-matrizes, constitui um processo de formação de palavras produtivo em português e, em muitos casos, é regido pelo princípio da analogia.

Assim, pela quarta proporcional, temos *milk + shake: milkshake :: McDonalds + milk-shake: x*, donde $x = Mcshake$.

O anúncio publicitário seguinte traz o neologismo *cãosumidor*, também um espelhamento analógico, em que, para a sua formação e interpretabilidade, foi evocada a unidade léxica *consumidor*.

Figura 4 – Anúncio publicitário



Fonte: Disponível em: https://www.instagram.com/p/B8AHPAojL0w/?utm_source=ig_web_copy_link. Acesso em: 20 abr. 2020.

Na criação do neologismo *cãosumidor* nota-se processo parecido, em que houve encurtamento apenas da segunda forma lexical (*cão + -sumidor*), caracterizando a chamada “combinação truncada”, como salienta Gonçalves (2016, p. 77).

Com isso, os neologismos *Mcshake* e *cãosumidor*, além da criatividade que revelam, têm sua força expressiva resultante da remissão formal às unidades léxicas *milk-shake* e *consumidor*, respectivamente, com a síntese de seus significados, em cuja operação está o princípio da analogia.

4. O PROCESSO DE DESAUTOMATIZAÇÃO NO GÊNERO PUBLICITÁRIO

Saussure (1978), ao tratar das relações sintagmáticas, estabeleceu as bases conceituais da fraseologia, disciplina linguística que posteriormente recebeu grande impulso pelos trabalhos de seu discípulo Charles Bally. A partir de então, as unidades fraseológicas têm recebido a atenção de variados estudiosos, tornando-se fecundo objeto de estudo.

As unidades fraseológicas nos textos publicitários servem para destacar as qualidades do produto, para influir na decisão do potencial consumidor e, sobretudo, para chamar a atenção. Zamora Muñoz (1999) destaca dois fatores que podem levar os publicitários a escolher fraseologismos para seus anúncios: o efeito das expressões sugestivas e a memorização do produto ou da marca.

Em consideração às lexias complexas, merecem atenção os graus de estruturação e automatização das construções sintáticas que, em escalas que vão do sintagma livre ao sintagma fixo, podem-se apresentar sob vários tipos de fraseologismos. Nesse aspecto, importa considerar as associações lexicais, de que resultam combinatórias estáveis com graus de coesão diversos. Para Saussure (1978, p. 209), tais unidades pertencem ao plano da língua e não ao plano da fala, por serem “loquções estereotipadas que não podem ser alteradas, ain-

da que se possa distinguir, pela reflexão, as suas partes significativas”. Trata-se de unidades do léxico formadas por mais de um elemento lexical, com elevado grau de fixidez em sua forma e em seu significado (Zuluaga, 1980). Contudo, a relação criatividade/produktividade está sempre presente nos comportamentos linguísticos dos usuários da língua, o que promove a inovação lexical. É nesse âmbito que atua a desautomatização, gerando formações novas decorrentes da alteração da forma e/ou do significado de uma estrutura matriz, anterior.

Um aspecto pelo qual a desautomatização pode ser observada é como um dispositivo de intertextualidade ou de polifonia textual (Grésillon; Maingueneau, 1984), uma vez que a manipulação de uma estrutura fixa provoca várias leituras possíveis do mesmo enunciado, bem como desencadeia todo um conjunto de relacionamentos e evocações.

Nesse contexto, as modificações são manipulações criativas que os falantes produzem, com certas finalidades. O resultado acaba sendo uma combinação léxica nova, não usual, decorrente da produção na reprodução. Ou seja, a combinatória modificada é uma unidade em parte produzida e em parte reproduzida pelo falante. É o resultado de um ato criativo e voluntário dentro dos limites de uma estrutura relativamente fixa e cristalizada (Zuluaga, 2007). É possível ilustrar tal fato considerando como estrutura matriz o nome do filme, um clássico bastante conhecido, a partir do qual a peça publicitária da Hortifruti apresenta a sua desautomatização, por meio de uma combinação léxica nova, que espelha a original, anterior.

Figura 5 – Anúncio publicitário



Fonte: Disponível em: <https://institucional.hortifruti.com.br/comunicacao/campanhas/hollywood/>. Acesso em: 23 jan. 2018.

Outro exemplo está no anúncio publicitário da figura 6, em que sobressai o jogo com as palavras pela alusão que faz à expressão fixa que designa a operação policial que se tornou famosa no Brasil por deter e prender diversos políticos e empresários brasileiros. Em tal anúncio, é formada uma expressão nova cujo sentido é construído não apenas em referência à nova embalagem do preservativo masculino, mas principalmente pela função do produto: “operação guarda-jato”.

Figura 6 – Anúncio publicitário



Fonte: Disponível em: <http://www.juruemdestaque.com/2017/05/preservativo-masculino-eros-premium-faz.html>. Acesso em: 20 ago. 2018.

5. PERSPECTIVAÇÃO PEDAGÓGICA

Dada a dinamicidade e a expansão do léxico, vários processos concorrem para a formação de palavras, muitos dos quais amplamente estudados nas gramáticas brasileiras, bem como em livros didáticos de português. Outros, no entanto, permanecem marginalizados por tais instrumentos didáticos, a despeito de sua produtividade, como os casos de analogia e desautomatização, que constituem o objeto de estudo neste capítulo.

Da análise desses processos, é possível alcançar uma visão ampliada da competência lexical do usuário geral da língua, especialmente quando se analisam os produtos deles decorrentes, na forma de neologismos, em manifestação em diversos gêneros discursivos.

Considerando a necessidade de se ter material teórico e material didático que possam ser úteis ao ensino do léxico, o estudo da neologia se reveste de importância pela possibilidade de lidar com o fenômeno do surgimento de palavras novas no português contemporâneo, acompanhando em tempo real a evolução da língua. Entretanto, estudar o léxico sob a perspectiva pedagógica implica em ter muitas informações sobre a palavra (Richards, 1976; 1985), tendo em consideração sua vasta tipologia de manifestação nos discursos oral e escrito.

Nesse contexto, ao se ter em conta o ensino do léxico do português, uma das principais dificuldades para apresentar coerentemente os processos lexicogenéticos no português, especialmente os processos de formação de palavras novas, é encontrar os critérios adequados nos quais se deve basear a apresentação desses fenômenos. Cumpre então se lembrar da importância do trabalho de Lewis (1993; 1997), em cujo modelo o léxico ocupa posição central, o chamado enfoque léxico (*The Lexical Approach*), no qual a unidade léxica é a unidade de aprendizagem.

Nesse contexto, é plenamente justificável o esforço por compreender, no que concerne à competência lexical, o que poderá informar um *corpus* de neologismos estabelecido. Neologismos que, no caso em pauta, são formações pelos processos de analogia e desautomatização, processos ainda marginalizados nas abordagens de muitas gramáticas tradicionais brasileiras.

O que se procura, nessa perspectiva, é relacionar a competência lexical, não com a fase de aquisição na infância, mas com o contínuo processo de desenvolvimento, que é o produto de aprendizagem por meio da educação sistemática, especialmente como se dá nos níveis avançados de educação.

Considerando a importância de se trabalhar, na sala de aula de língua portuguesa, com a diversidade de gêneros textuais, o anúncio publicitário está entre aqueles ainda pouco explorados, pelo que se observa em certas análises de livros didáticos de português (Cruz, 2015).

Propomos, em face disso, dar ênfase à utilização do texto publicitário nas aulas de língua portuguesa, pondo em destaque todo o seu arranjo criativo, estruturante do gênero, tal como se observa no texto literário, em sua criação estética. Ademais, convém ressaltar a relevância do uso de materiais didáticos autênticos no âmbito da sala de aula, além dos textos didáticos já criados ou organizados com tal finalidade, como bem atestou Martins (2016).

CONSIDERAÇÕES FINAIS

Com tais produtos neológicos, expostos nas seções 3 e 4, assume-se que analogia e desautomatização são processos convergentes de criação neológica; e ressalta-se a relevância de um tratamento pedagógico para tais casos em face da precarização do ensino do léxico

(Antunes, 2012, p. 13), em especial do estudo sobre neologia, pelo que se observa nas gramáticas escolares brasileiras. Assim, a existência de sérias lacunas nos estudos linguísticos sobre o português brasileiro, especialmente no que concerne ao léxico, constitui forte motivação para se empreender uma pesquisa sobre a produtividade lexical, aqui delimitada ao estudo específico dos processos neológicos de analogia e desautomatização no gênero publicitário.

Enquanto o princípio da analogia é usado inconscientemente pelo usuário da língua, a desautomatização é o resultado de manipulações conscientes, projetadas para os propósitos persuasivos deste mesmo usuário, o que torna a desautomatização dependente da sua criatividade e do efeito que este procura provocar no interlocutor.

Com isso, se a analogia favorece a inovação lexical pelo espelhamento de formas anteriores, também a desautomatização de combinações fixas gera produtos neológicos que se espelham em estruturas anteriores.

Ocorre a desautomatização de combinações fixas porque existe uma memória coletiva fraseológica, isto é, esse fenômeno não seria tão recorrente se os usuários da língua não possuísem um conhecimento prévio e compartilhado do valor semântico e cultural do fraseologismo. Neste caso, a ativação da memória deve-se ao fato de a formação nova espelhar-se na estrutura original. Ou seja, uma combinatória léxica, estrutura fixa e institucionalizada, presente no léxico mental do usuário da língua, quando sofre manipulação que lhe altera a forma é posteriormente por este reconhecida, graças à remissão que faz à forma matriz, anterior.

Na mesma linha de ação, o princípio da analogia, subjacente a processos de formação de palavras (Basílio, 1997; 2010), converge

para o mesmo ponto que faculta a desautomatização de combinações fixas, por meio da formação de unidades fraseológicas novas.

Em consideração a tudo isso, ressaltamos a devida importância ao desenvolvimento da competência lexical que, tendo em pauta o estudo da neologia, reserva importante papel para o gênero publicitário, como instrumento didático autêntico.

O esforço em preencher lacunas existentes, no âmbito das gramáticas escolares e dos livros didáticos de português, com o estudo da analogia e desautomatização como processos neológicos, justifica-se principalmente para os estudantes do ensino médio (Martins, 2016).

Em últimas palavras, a didatização de textos publicitários em sala de aula de língua portuguesa é um importante recurso para se trabalhar a palavra que, entre outros aspectos, pode-se destacar por seu caráter neológico, apresentando novidade semântico-lexical.

REFERÊNCIAS

ANTUNES, Irandé. **O território das palavras**: estudo do léxico em sala de aula. São Paulo: Parábola Editorial, 2012.

BAKHTIN, Mikhail. **Estética da criação verbal**. Trad. Paulo Bezerra. São Paulo: Martins Fontes, 2003.

BASILIO, Margarida. O princípio da analogia na constituição do léxico: regras são clichês lexicais. **Veredas**: revista de estudos linguísticos, Juiz de Fora, v. 1, n. 1, p. 1-14, 1997.

BASILIO, Margarida. Abordagem gerativa e abordagem cognitiva na formação de palavras: considerações preliminares. **Revista Linguística**, Rio de Janeiro, v. 6, n. 2, dez. 2010. Disponível em: <https://revistas.ufrj.br/index.php/rl/article/view/4444>. Acesso em: 20 ago. 2015.

BAUER, Laurie. **English Word-formation**. Cambridge: Cambridge University Press, 1983.

CRUZ, Ana Flávia Torquetti. **Léxico e livro didático: estratégias para o desenvolvimento da leitura.** 2015. 134 f. Dissertação (Mestrado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2015.

FERRAZ, Aderlande Pereira. Do observatório de neologia para a sala de aula: contribuição para o ensino do léxico. *In: CARDOSO, Elis de Almeida et al. (Orgs.). Os estudos lexicais em diferentes perspectivas.* São Paulo: FFLCH/USP, 2020. v. 8. p. 165-179.

FERRAZ, Aderlande Pereira. O discurso publicitário e a criação de palavras novas: o desenvolvimento da competência lexical. **Fólio – Revista de Letras,** Vitória da Conquista, v. 11, n. 2, p. 49-69, jul./dez. 2019.

FERRAZ, Aderlande Pereira. Publicidade: a linguagem da inovação lexical. *In: ALVES, Ieda Maria (Org.). Neologia e neologismos em diferentes perspectivas.* São Paulo: Paulistana, 2010. p. 251-275.

FERRAZ, Aderlande Pereira. Os neologismos no desenvolvimento da competência lexical. *In: HENRIQUES, Claudio Cezar; SIMÕES, Darcília (Orgs.). Língua portuguesa, educação & mudança.* Rio de Janeiro: Europa, 2008. p. 146-162.

GONÇALVES, Carlos Alexandre. **Atuais tendências em formação de palavras.** São Paulo: Contexto, 2016.

GRÉSILLON, Almuth; MAINGUENEAU, Dominique. Polyphonie, proverbe et détournement ou un proverbe peut en cacher un autre. **Langages,** Paris, n. 73, p. 112-185, 1984.

GUILBERT, Louis. **La créativité lexicale.** Paris: Larousse, 1975.

LEWIS, Michael. **The lexical approach.** London: Language Teaching Publications, 1993.

LEWIS, Michael. **Implementing the lexical approach.** London: Language Teaching Publications, 1997.

MARTINS, Élide Ferreira. Neologismos semânticos e texto publicitário: explorando o léxico do português em sala de aula. *In: FERRAZ, Aderlande Pereira (Org.). O léxico do português em estudo na sala de aula.* Araraquara: Letraria, 2016. p. 31-52.

POTTIER, Bernard *et al.* **Estruturas linguísticas do português.** São Paulo: Difusão Européia do Livro, 1972.

RICHARDS, Jack. The role of vocabulary teaching. **TESOL Quarterly**, New York, v. 10, n. 1, p. 77-89, mar. 1976.

RICHARDS, Jack. Lexical knowledge and the teaching of vocabulary. *In*: RICHARDS, Jack. **The context of language teaching**. Cambridge: Cambridge University Press, 1985. p. 176-188.

SABLAYROLLES, Jean-François. Le sentiment néologique. *In*: SABLAYROLLES, Jean-François (Org.). **L'innovation lexicale**. Paris: Honoré Champion Éditeur, 2003. p. 279-295.

SANDMANN, Antônio José. **Competência lexical**: produtividade, restrições e bloqueio. Curitiba: UFPR, 1991.

SAUSSURE, Ferdinand de. **Curso de linguística geral**. 4. ed. Tradução: José Victor Aragão. Lisboa: Publicações Dom Quixote, 1978.

ZAMORA MUÑOZ, Pablo. Desautomatización y traducción de expresiones fijas italianas en español. **Cahiers du PROHEMIO**, Orléans, n. 3, p. 425-440, 1999.

ZULUAGA, Alberto. **Introducción al estudio de las expresiones fijas**. Tübingen: Max Hueber, Verlag, 1980.

ZULUAGA, Alberto. Análisis y traducción de unidades fraseológicas desautomatizadas. **PhiN**, v. 16, p. 67-83, 2007. Disponível em: <http://web.fu-berlin.de/phn/phn16/p16t5.htm>. Acesso em: 11 nov. 2014.

TOPONÍMIA EM LIBRAS: ICONICIDADE E EMPRÉSTIMOS NA FORMAÇÃO DOS SINAIS

Alexandre Melo de Sousa

CONSIDERAÇÕES INICIAIS

A Língua Brasileira de Sinais (Libras) é a língua dos surdos brasileiros que vivem, espacialmente, nos centros urbanos (Quadros, 2019). O reconhecimento legal da Libras se deu a partir da Lei 10.436/2002, que a oficializou como língua nacional de comunicação e expressão dos surdos brasileiros, e foi regulamentada pelo Decreto 5626/2005. Esse marco legal foi importante para a criação de disciplinas de Libras nos cursos de Licenciatura, para a criação dos cursos de Letras Libras, para o estabelecimento de políticas linguísticas, educacionais e de acessibilidade entre outros avanços que impulsionaram as pesquisas com foco na Libras, especialmente no âmbito das pós-graduações.

Em 2021, a Lei 14.191/2021 incluiu a educação bilíngue de surdos na LDB (Lei de Diretrizes e Bases da Educação Nacional, Lei 9.394/1996) como uma modalidade de ensino, desvinculada do ensino especial. Esse marco legal constitui o reconhecimento da diferença surda e dos papéis que a Libras e a língua portuguesa escrita têm na perspectiva interacional e educacional da pessoa surda.

A Língua Brasileira de Sinais se distingue de outras línguas orais existentes no país pela sua modalidade de produção que utiliza o espaço e os articuladores corporais (mãos, braços, face etc.) e de percepção, que acontece através da visão, para estabelecer seu processo de interação a partir de construções (corp)orais. A Libras também se distingue de outras línguas de sinais minoritárias do país, que emerge e se consolida entre surdos de comunidades indígenas ou de comunidades isoladas. Como as demais línguas, a Libras possui todos os níveis de análise linguística: fonológico, morfológico, sintático, semântico, discurso, lexical etc. Neste texto, nos interessa o léxico toponímico em Libras.

O objetivo deste estudo é mostrar como a iconicidade e o português escrito influenciam a produção dos topônimos em Libras. Para isso, definiremos o que é toponímia e quais as especificidades da toponímia em línguas de sinais; em seguida, trataremos a respeito da expansão e consolidação dos estudos toponímicos em Libras no Brasil, destacando desde os primeiros trabalhos sobre o tema até as pesquisas mais atuais. Na seção seguinte, nos deteremos na iconicidade dos topônimos em Libras e, logo após, falaremos da influência do português escrito na formação dos sinais que nomeiam espaços geográficos em Libras. Para fechar o texto, faremos algumas considerações conclusivas.

1. TOPONÍMIA E TOPONÍMIA EM LIBRAS

Como uma das subáreas da Onomástica – disciplina linguística que estuda os nomes próprios em geral – a Toponímia dedica-se especificamente ao estudo dos nomes próprios de lugares (Dick, 1990; 1992; Isquerdo, 2019). Diferentemente do que acontecia nos primeiros estudos toponímicos brasileiros – Sampaio (1901); Oliveira

(1957); Cardoso (1961); Drumond (1965); Mello (1967) – a partir da década de 1980, os estudos dos nomes próprios de lugares extrapolaram o viés etimológico de base (quase que exclusivamente) indigenista, e tomaram o topônimo em sua totalidade estrutural e semântico-motivacional. Esse novo olhar da pesquisa toponímica deve-se aos estudos de Maria Vicentina do Amaral Dick, da Universidade de São Paulo, publicados em Dick (1990; 1992; 1996; 1998; 1999), entre outros (Sousa; Dargel, 2017; 2020).

Poderíamos citar inúmeras contribuições das pesquisas de Dick, mas destacamos, aqui, o modelo taxionômico proposto pela autora, em que os nomes de lugares são distribuídos de acordo com seus referentes motivacionais (ou seja, as possíveis causas denominativas) em dois grandes grupos: os de natureza física (onze categorias: *astrotopônimos*, *cromotopônimos*, *fitotopônimos*, *litotopônimos*, *zootopônimos* etc.) e os de natureza antropocultural (dezesseis categorias: *antropotopônimos*, *etnotopônimos*, *historiotopônimos* etc.). Esse modelo classificatório tem servido de base teórica para a grande maioria de pesquisas toponímicas desenvolvidas no Brasil e, especialmente, para os projetos de Atlas toponímicos já desenvolvidos e em desenvolvimento nas universidades brasileiras: *Projeto Atlas Toponímico do Paraná* (ATEPAR), *Projeto Atlas Toponímico do Estado de Mato Grosso do Sul* (ATEMS), *Projeto Atlas Toponímico do Estado de Minas Gerais* (ATEMIG), *Projeto Atlas Toponímico do Tocantins* (ATT), *Projeto Atlas Toponímico de Origem Indígena do Tocantins* (ATITO), *Projeto Atlas Toponímico do Estado do Maranhão* (ATEMA), *Projeto Atlas Toponímico da Bahia* (ATOBAH), *Projeto Atlas Toponímico da Amazônia Ocidental Brasileira* (ATAOB), para citar alguns.

O Projeto ATAOB possui um diferencial em relação aos demais Atlas, pois inclui os dados toponímicos em Libras do estado do Acre. A inclusão dos dados em língua de sinais possibilitou obser-

var, contrastivamente em relação às línguas orais, os componentes estruturais dos sinais toponímicos e a base semântico-motivacional inerente aos topônimos em Libras (Sousa, 2019; 2022; 2023).

2. EXPANSÃO E CONSOLIDAÇÃO DOS ESTUDOS SOBRE TOPONÍMIA EM LIBRAS

Os estudos toponímicos em línguas de sinais são muito recentes no Brasil. Os primeiros datam da primeira década do segundo milênio, com as pesquisas de Souza-Junior (2012) e Aguiar (2012). Essas pesquisas descrevem os topônimos em Língua Brasileira de Sinais (Libras) de formas distintas: enquanto o primeiro utiliza o viés proposto por Dick (1990), tanto no que se refere ao aspecto estrutural, quanto ao aspecto semântico-motivacional, o segundo aborda as questões relacionadas à iconicidade e aos empréstimos da língua oral na formação dos sinais que nomeiam os espaços selecionados no dicionário *DeitLibras*, de Capovilla e Raphael (2009).

No Acre, os dados toponímicos em Libras passaram a ser incorporados ao *Atlas Toponímico da Amazônia Ocidental Brasileira* (Projeto ATAQB), sediado na Universidade Federal do Acre, a partir da pesquisa de Bezerra (2015), que analisou as estruturas fonológica e morfológica dos sinais dos municípios acreanos. De lá para cá, outros estudos foram desenvolvidos como o de Bezerra (2016), que discutiu a questão da iconicidade em sinais de quatro cidades que formam a Regional do Alto Acre (Assis Brasil, Brasiléia, Epitaciolândia e Xapuri), e o de Alemão (2017), que deu início à elaboração da ficha lexicográfico-toponímica em versão digital, para o armazenamento dos dados em Libras no ATAQB.

A pesquisa dos topônimos em Libras se expandiu de tal modo que, a partir de 2018, os dados em Libras constantes no projeto *Atlas Toponímico da Amazônia Ocidental Brasileira* migraram para o projeto *Toponímia em Libras*, sediado na mesma instituição. A pesquisa de Sousa (2018), que discute as metodologias para as análises dos topônimos em línguas de sinais, considerando as especificidades das línguas de modalidade visual-espacial, dá início ao novo ciclo de investigações toponímicas em Libras no estado do Acre. Dados relacionados à configuração das fichas lexicográfico-toponímicas digitais foram discutidos em Sousa e Quadros (2019c); a construção de um *Web Software* para disponibilizar os dados toponímicos dos municípios acreanos foi descrita em Sousa e Quadros (2019a; 2019b) e a proposta de análise estrutural dos topônimos em Libras foi apresentada em Sousa (2019; 2022; 2023).

Merecem destaque outras pesquisas que se debruçaram sobre os estudos toponímicos em Libras: Jesus (2019), Ferreira (2019), Miranda (2020), Marcelino (2021), Carmo (2021) e Souza (2023). Os dois primeiros trabalhos estão vinculados ao Projeto Estudo bilíngue da toponímia de Feira de Santana-BA: Português-Libras, sediado na Universidade Estadual de Feira de Santana. Jesus (2019) estuda a motivação dos topônimos em Libras dos sinais que nomeiam os bairros de Feira de Santana, enquanto Ferreira (2019) dedica-se ao estudo dos fatores que motivaram as criações dos sinais em Libras que dão nome às ruas e avenidas do centro comercial da mesma cidade baiana.

Miranda (2020), por sua vez, debruça-se sobre o processo de criação dos sinais que nomeiam as cidades de Tocantins. A pesquisadora, além de estudar as propriedades articulatórias e de origem dos sinais toponímicos em Libras, busca identificar fatores motivacionais para a criação dos topônimos selecionados.

Já Urbanski, Ferreira e Xavier (2020) analisaram oitocentos e oitenta e oito topônimos da Libras que nomeiam cidades de quatorze estados do Brasil e de algumas regiões administrativas do Distrito Federal. Os dados apontaram grande influência do português nos sinais toponímicos.

Marcelino (2021), por seu turno, analisou os sinais das cidades acreanas (vinte e dois sinais) quanto à variação fonológica, morfológica e lexical. Os dados foram extraídos de vídeos disponíveis no YouTube (canal da Profa. Ianele Viviane Vital Pereira de Melo, da Universidade Federal do Acre; do Centro de Apoio ao Surdo do Acre – CAS/AC; e do *web software* Toponímia em Libras).

Carmo (2021) estudou os sinais em Libras que dão nome a doze espaços públicos de lazer de Rio Branco, capital do Acre. Cada sinal foi analisado quanto à estrutura fonomorfológica, quanto aos aspectos motivacionais – tomando como base os estudos de Dick (1990) e Sousa (2019) – e quanto à iconicidade dos sinais selecionados.

Por fim, Souza (2023) registrou e analisou cento e doze sinais toponímicos que nomeiam cidades de Goiás, procurando evidências e influências linguísticas e extralinguísticas no processo de nomeação. O estudo apresenta interação entre a Toponímia, a Libras e a Linguística de *Corpus*.

De um modo geral, essas pesquisas tratam da formação estrutural do topônimo em Libras e das questões motivacionais de criação dos sinais que nomeiam os espaços geográficos em Libras. Contudo, nem todas seguem apenas a proposta de Dick (1990) para a análise dos dados. Neste texto, por exemplo, trataremos das questões relacionadas à iconicidade e à influência do português escrito na criação dos sinais dos municípios acreanos, em Libras.

3. A ICONICIDADE NOS SINAIS TOPONÍMICOS

A iconicidade é definida nos dicionários de Linguística como a relação direta que há entre a forma linguística e seu referente (Trask, 2004; Dubois *et al.*, 1973). Polguère (2018, p. 36), contudo, defende a existência de uma “relação de analogia” entre o conteúdo e a forma do signo.

Em línguas de sinais, a iconicidade é um traço partícipe da estrutura da língua (Perniss, 2007), especialmente quando se observa que o modo de produção dos sinais, envolvendo articuladores (corporais, como mãos, braços etc.), em movimento, num dado espaço tridimensional, conduz a relação entre a construção sinalizada e seu referente (Taub, 2001; Perniss, 2007; Lillo-Martin; Gajewskay, 2014; Quadros, 2019).

Para Lillo-Martin e Gajewskay (2014), a iconicidade tem uma função fundamental no processamento lexical (morfológico e semântico), no esboço de estruturas conceituais abstratas, no estabelecimento das referências espaciais, pronominais etc. Do mesmo modo, Quadros (2019) ressalta que a iconicidade tende a motivar a formação de vários sinais.

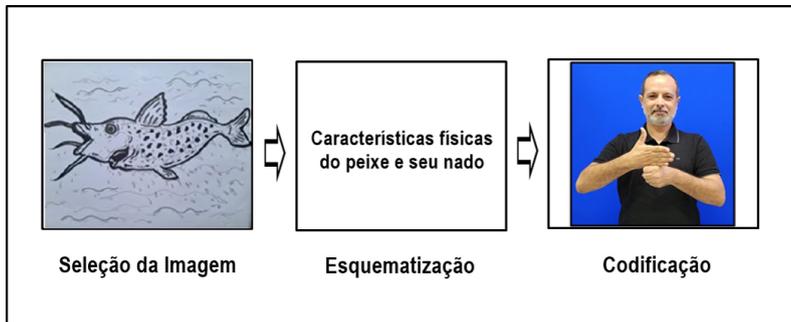
Para Carneiro (2016, p. 110), a iconicidade, inerente às línguas de sinais, “possibilita ao sinalizador codificar concepções diversas e construir estruturas icônicas, capazes de transmitir grande número de informações de maneira simultânea”.

Taub (2001) descreve o processo de formação dos itens lexicais icônicos, com base no processo que envolve três etapas: a seleção, a esquematização e a codificação. A **etapa da seleção** corresponde à escolha de uma imagem representativa e coerente com o referente; a **etapa da esquematização** corresponde à reformulação da imagem

representativa, em seus aspectos mais relevantes, que possam se estruturar na articulação fonética e na construção semântica da língua; a **etapa da codificação** corresponde à materialização da forma linguística – o sinal propriamente dito.

Com base na proposta de Taub (2001), demonstraremos o processo de criação dos sinais toponímicos BUJARI, BRASILÉIA, MÂNCIO LIMA, RIO BRANCO, SENA MADUREIRA e TARAUCÁ.

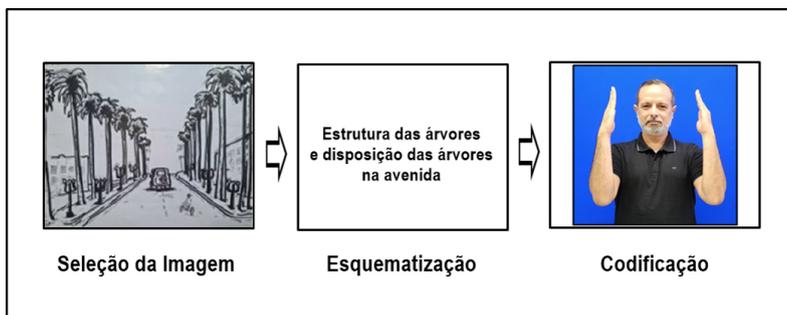
Figura 1 – Iconicidade no sinal toponímico BUJARI



Fonte: adaptado de Sousa (2022).

Para a formação do sinal toponímico BUJARI, selecionou-se como imagem referente ao peixe Mandi (bagre), característico da região. Para a esquematização, consideraram-se as características físicas do peixe (o formato, em especial) e a ação de nadar. Na codificação do sinal, o formato físico do peixe foi projetado na configuração de mão, e o nado, foi refletido no movimento do sinal.

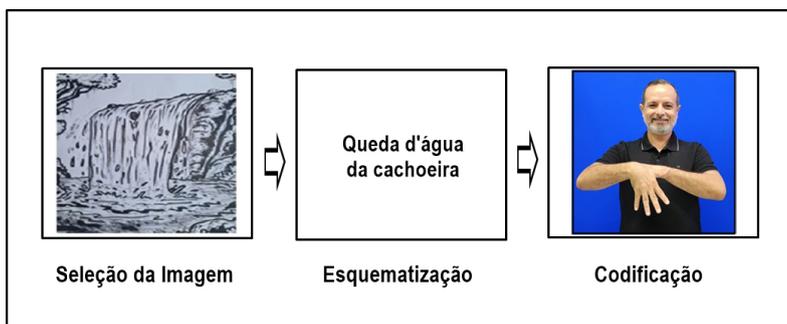
Figura 2 – Iconicidade no sinal toponímico BRASILEIA



Fonte: adaptado de Sousa (2022).

Para a produção do sinal toponímico BRASILEIA, selecionou-se a imagem das árvores, localizadas na principal rua da cidade, organizadas em fileiras. Para a etapa da esquematização, consideraram-se a estrutura física das árvores e o modo como elas estão dispostas ao longo da rua. Na etapa de codificação do sinal, projetou-se a forma física das árvores na configuração das mãos e dos braços. A organização das árvores, ao longo da rua, foi refletida no movimento sincronizado dos braços para a frente do corpo no sinalizante, em espaço neutro.

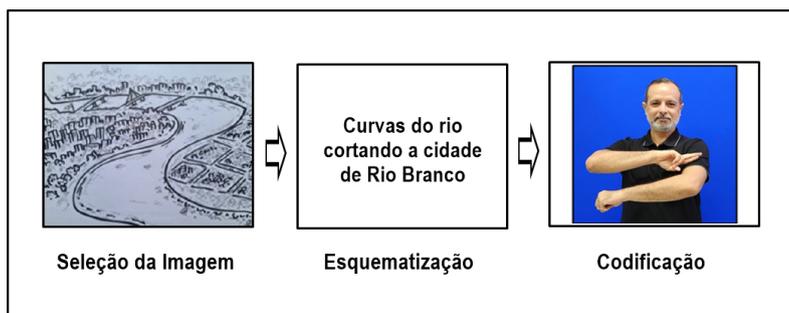
Figura 3 – Iconicidade no sinal toponímico MÂNCIO LIMA



Fonte: adaptado de Sousa (2022).

Para a produção do sinal MÂNCIO LIMA selecionou-se a imagem da cachoeira localizada na região do espaço nomeado. Para a esquematização, considerou-se a queda d'água. Para a codificação do sinal, a mão ativa projeta uma queda iniciando no dorso da mão passiva em direção ao chão, de modo que a configuração da mão faz referência à água, e o movimento dos dedos e da mão, em queda em direção ao chão, refletem a água caindo. Vale lembrar que o sinal MÂNCIO LIMA inicia com a configuração da mão ativa em “M”, em referência à letra inicial do topônimo em língua portuguesa, tocando o dorso da mão passiva.

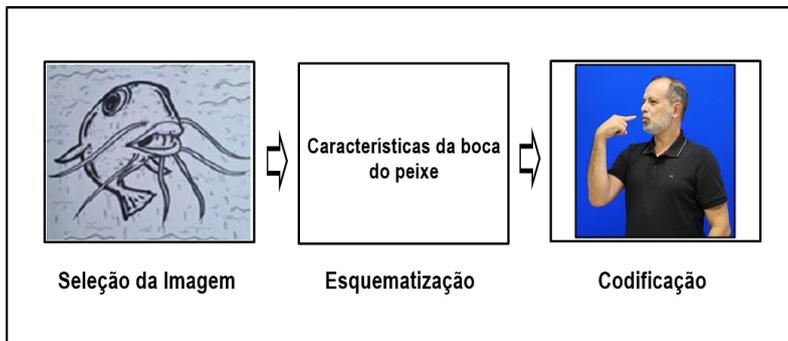
Figura 4 – Iconicidade no sinal toponímico RIO BRANCO



Fonte: adaptado de Sousa (2022).

Para a produção do sinal toponímico RIO BRANCO, selecionou-se a imagem do Rio Acre, localizado na capital acreana. Para a esquematização, consideraram-se as características físicas do rio, de modo especial, suas curvas sinuosas. Na codificação do sinal, a mão ativa assume duas configurações diferentes: inicialmente, a configuração fica em “R” e, posteriormente, em “B”. A mão ativa percorre o braço, desde o ponto mais próximo ao ombro até o antebraço, finalizando próximo à mão. Durante esse percurso, a mão ativa faz uma curva no cotovelo, em referência à curva do rio.

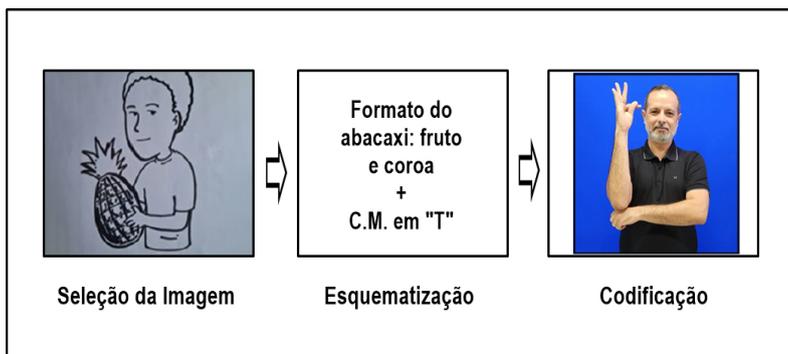
Figura 5 – Iconicidade no sinal toponímico SENA MADUREIRA



Fonte: adaptado de Sousa (2022).

Para a formação do sinal SENA MADUREIRA selecionou-se a imagem do peixe Mandi (bagre). No entanto, para a esquematização, o elemento destacado foi a boca arredondada e saliente do peixe. Na codificação, o articulador ativo assume a seguinte configuração de mão: a mão fica cerrada e o dedo indicador estendido aponta para os lábios (levemente projetados para a frente) fazendo movimentos circulares em torno dos lábios do sinalizante. Trata-se de uma referência à boca saliente do peixe, como dissemos anteriormente.

Figura 6 – Iconicidade no sinal toponímico TARAUCÁ



Fonte: adaptado de Sousa (2022).

Para a formação do sinal toponímico TARAUCÁ selecionou-se a imagem do abacaxi – fruto característico da cidade. Para a etapa de esquematização, consideraram-se as características físicas do abacaxi: o fruto e a coroa espinhosa. Para a codificação, projetaram-se as características do abacaxi nos articuladores ativos: o braço estendido e apoiado na mão passiva, e a mão ativa com configuração em “T”, com os demais dedos abertos refletem a coroa do abacaxi. O movimento circular dos articuladores ativos faz referência ao formato da fruta.

4. A INFLUÊNCIA DO PORTUGUÊS NOS SINAIS TOPONÍMICOS

Em Sousa (2019; 2022; 2023), discutimos sobre os tipos de formação dos sinais toponímicos em Libras e os classificamos em quatro tipos: simples, simples híbrido, composto e composto híbrido – numa ampliação da proposta de Dick (1990), considerando as especificidades das línguas de modalidade visual-espacial, como a Libras. Essa proposta pode ser resumida assim: a) **formação simples**, quando há apenas um formante na língua de sinais nativa (ou seja, em Libras); b) **formação simples híbrida**, quando há apenas um formante com empréstimo da língua oral em sua estrutura; c) **formação composta**, quando há mais de um formante, e todos os elementos são da língua de sinais nativa e, d) **formação composta híbrida**, quando contém mais de um formante: sendo pelo menos um da língua de sinais nativa, e em algum deles haja empréstimo de língua oral ou outra língua de sinais distinta da nativa. Vejamos os dois sinais a seguir:

Figura 7 – Formações toponímicas simples



Fonte: elaborado pelo autor

Pela ilustração é possível verificar que ambos os sinais possuem a mesma configuração de mão passiva (neste caso, a mão esquerda). Contudo, o primeiro sinal, BUJARI, é realizado com a mão aberta e os dedos juntos fazendo referência ao peixe, como mostramos na seção anterior. O sinal toponímico em questão possui um único formante e ele é da língua nativa. Trata-se, portanto, de um sinal simples.

O sinal CAPIXABA, contudo, apresenta configuração de mão em “C”, fazendo clara referência ao nome da cidade em língua portuguesa. Nesse caso, embora possua também um único formante, há empréstimo da língua oral em sua formação. Trata-se de um sinal simples híbrido. Agora, vejamos os sinais compostos a seguir:

Figura 8 – Formações toponímicas compostas



Fonte: elaborado pelo autor a partir de Sousa (2019) e Carmo (2021)

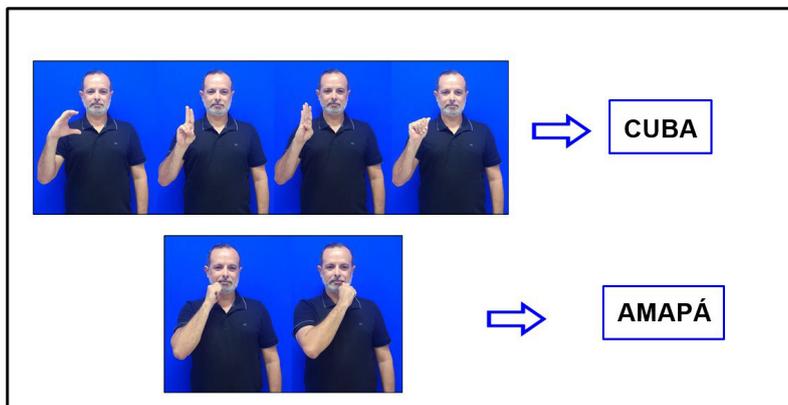
Nos casos ilustrados na figura 8, os sinais possuem dois formantes cada. No primeiro caso, CALÇADÃO DA GAMELEIRA, o sinal é formado por BANDEIRA + ESCADARIA – dois sinais em língua nativa. Trata-se, nesse caso, de um sinal composto.

No segundo caso, MÂNCIO LIMA, o sinal possui dois formantes: o primeiro apresenta a configuração em M, fazendo referência à letra inicial do nome da cidade em português. Em seguida, temos o sinal CACHOEIRA, fazendo referência ao elemento hidrográfico situado no local – como mostramos na seção anterior. Trata-se de um sinal composto híbrido.

Como vemos, alguns sinais toponímicos podem ser formados por empréstimos da língua oral. Faria-Nascimento (2009) classifica esse processo como empréstimo por transliteração, que pode ser dos tipos pragmática (quando o empréstimo é datilológico, ou seja, soletrado a partir das configurações de mão) e lexicalizada (quando o sinal possui apenas algumas configurações com referência às letras da palavra em língua oral). Nesse último caso, podemos verificar a transliteração da letra inicial, ou como chamou Ferreira-Brito (1995): empréstimo por inicialização.

No caso dos topônimos em Libras, alguns exemplos podem ser observados a seguir.

Figura 9 – Transliteração



Fonte: elaborado pelo autor a partir de Aguiar (2012) e Sousa (2023)

No primeiro caso temos o topônimo CUBA realizado de forma datilológica, a partir da realização “soletrada” de cada configuração correspondente às letras que formam a palavra em língua portuguesa: C + U + B + A. Trata-se de uma transliteração pragmática, como classificada por Farias-Nascimento (2009).

No segundo caso, o topônimo em Libras AMAPÁ é formado pela configuração de mão em “A”, tocando nas extremidades dos lábios do sinalizante. A configuração faz referência às letras inicial e final do topônimo em língua oral. Temos, nesse caso, uma transliteração lexicalizada.

Em Sousa (2019), observamos que alguns topônimos podem apresentar, em suas formações, empréstimos de outros sinais gráficos da língua oral, como números, arroba etc. Vejamos o exemplo a seguir:

Figura 10 – Sinal BAIRRO QUINZE



Fonte: Sousa (2023)

Nesse caso, chamamos o processo de empréstimo por *transemiotização*, numa ampliação da proposta de Farias-Nascimento (2009) e que tem grande produtividade nos topônimos em Libras. Em Rio Branco há a rua Marechal Deodoro, no Centro, onde está localizado o banco Itaú, cujo sinal faz referência ao sinal gráfico arroba “@”.

Quando pensamos na classificação taxionômica dos topônimos em Libras que fazem referência às letras dos topônimos em língua oral, adotamos a classificação *grafotopônimo*, de Souza-Júnior (2012), para aqueles que possuem datilologia completa do topônimo correspondente em português (ou outra língua oral), como no caso de CUBA. Se o topônimo em Libras for formado com configuração de apenas algumas letras (inicial, inicial e final etc.) do topônimo correspondente em língua oral, classificamos como *acronimotopônimo*, segundo Francisquini (1998), por fazerem referência às siglas e abreviações.

É válido ressaltar que, quando a configuração de mão referente à(s) letra(s) do topônimo em língua oral estiver incorporada na estru-

tura do sinal, como no sinal TARAUCÁ (figura 6), prevalece, para fins de classificação taxionômica, o referente icônico do sinal, no caso, o fruto abacaxi. Assim, temos um *fitotopônimo*.

CONSIDERAÇÕES FINAIS

Os aspectos culturais sempre foram destacados, nos diversos estudos toponímicos, como essenciais no ato de nomeação de um determinado espaço geográfico. Na toponímia em Libras, a cultura visual dos nomeadores, os surdos, é refletida na constituição dos sinais, de modo notório quando são observadas a iconicidade dos sinais. A forma do topônimo em Libras, como mostramos em alguns exemplos, fazem referência direta aos elementos que motivaram a criação do sinal.

Do mesmo modo, percebemos o quanto a língua portuguesa escrita influencia a criação dos sinais toponímicos. Possivelmente, o fato de as línguas (Libras e Português) conviverem socialmente no mesmo espaço e o sujeito surdo fazer uso das duas línguas (no caso do português, na modalidade escrita), seja um fator influenciador da presença de configurações de mão representantes das letras do topônimo em Português – os nomes das cidades, dos bairros, das ruas, das praças estão escritos nos diversos espaços para a identificação urbana.

Os estudos toponímicos em Libras têm avançado. Embora, como afirmamos anteriormente, os pesquisadores tenham adotado metodologias e formas classificatórias diferentes, entendemos que esse é o caminho natural para uma área nova nos estudos linguísticos aplicados a Libras e que revela importância para a consolidação das pesquisas.

REFERÊNCIAS

AGUIAR, Mônica Cruz de. Descrição e análise dos sinais topônimos em Libras. In: ALBRES, Neiva de Aquino; XAVIER, André Nogueira (Orgs.). **Libras em estudo: descrição e análise**. São Paulo: FENEIS, 2012. p. 109-121.

ALEMÃO, Vicharlisson Brito. Projeto ATAQB: armazenamento e dados em Libras. **Relatório Final Programa Institucional de Bolsas de Iniciação Científica (PIBIC)**. Universidade Federal do Acre. Rio Branco: UFAC, 2017.

BEZERRA, Manuella Trindade. Formação dos sinais toponímicos acreanos. **Relatório Final Programa Institucional de Bolsas de Iniciação Científica (PIBIC)**. Universidade Federal do Acre. Rio Branco: UFAC, 2015.

BEZERRA, Manuella Trindade. Sinais toponímicos do Acre: a iconicidade no processo de formação. **Relatório Final Programa Institucional de Bolsas de Iniciação Científica (PIBIC)**. Universidade Federal do Acre. Rio Branco: UFAC, 2016.

BRASIL. Lei nº 9.394, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional. **Diário Oficial da União**, Brasília, DF, 23 dez. de 1996. Disponível em: https://www.planalto.gov.br/ccivil_03/LEIS/L9394.htm. Acesso em: 02 out. 2021.

BRASIL. Decreto Nº 5.626, de 22 de dezembro de 2005. Regulamenta a Lei nº 10.436, de 24 de abril de 2002, que dispõe sobre a Língua Brasileira de Sinais - Libras, e o art. 18 da Lei nº 10.098, de 19 de dezembro de 2000. **Diário Oficial da União**, Brasília, DF, 23 dez. de 2005. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2005/decreto/d5626.htm. Acesso em: 02 out. 2021.

BRASIL. Lei nº 10.436, de 24 de abril de 2002. Dispõe sobre a Língua Brasileira de Sinais - Libras e dá outras providências. **Diário Oficial da União**, Brasília, DF, 25 abr. de 2002. Disponível em: http://www.planalto.gov.br/ccivil_03/LEIS/2002/L10436.htm. Acesso em: 02 out. 2021.

BRASIL. Lei nº Lei 14.191/2021, de 3 de agosto de 2021. Altera a Lei nº 9.394, de 20 de dezembro de 1996 (Lei de Diretrizes e Bases da Educação Nacional), para dispor sobre a modalidade de educação bilíngue de surdos. **Diário Oficial da União**, Brasília, DF, 4 ago. de 2021. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/lei/l14191.htm. Acesso em: 02 out. 2021.

CAPOVILLA, Fernando César; RAPHAEL, Walkíria Duarte. **Dicionário enciclopédico ilustrado trilingue da língua de sinais brasileira**. São Paulo: Edusp, 2009.

CARDOSO, Armando Levy. **Toponímia brasílica**. Rio de Janeiro: Biblioteca do Exército, 1961.

CARMO, Felipe dos Santos do. **Toponímia em Libras dos parques, praças e espaços de lazer em Rio Branco (AC): análise dos aspectos formais e motivacionais dos sinais que nomeiam os espaços urbanos**. 2021. Trabalho de Conclusão de Curso (Graduação em Letras Libras) – Universidade Federal do Acre, Rio Branco, 2021.

CARNEIRO, Bruno Gonçalves. Ampliação lexical da Libras: aspectos icônicos. **Leitura**, v. 1, n. 57, p. 104-119, 2016. Disponível em: <https://www.seer.ufal.br/index.php/revistaleitura/article/view/2840>. Acesso em: 11 out. 2021.

DICK, Maria Vicentina de Paula do Amaral. **A motivação toponímica e a realidade brasileira**. São Paulo: Arquivo do Estado, 1990.

DICK, Maria Vicentina de Paula do Amaral. **Toponímia e antroponímia do Brasil: coletânea de estudos**. 3. ed. São Paulo: Serviço de Artes Gráficas da FFLCH/USP, 1992.

DICK, Maria Vicentina de Paula do Amaral. Atlas toponímico: um estudo de caso. **Acta Semiotica et Lingvistica**, São Paulo, v. 6, p. 27-44, 1996.

DICK, Maria Vicentina de Paula do Amaral. Atlas toponímico: um estudo dialetológico. **Revista Philologus**, Rio de Janeiro, v. 10, p. 61-69, 1998.

DICK, Maria Vicentina de Paula do Amaral. **A dinâmica dos nomes na Cidade de São Paulo 1554-1897**. São Paulo: ANNABLUME, 1997.

DRUMOND, Carlos. **Contribuições do bororo à toponímia brasílica**. São Paulo: USP/IEB, 1965.

DUBOIS, Jean *et al.* **Dicionário de Linguística**. São Paulo: Cultrix, 1973.

FARIA-NASCIMENTO, Sandra Patrícia. **Representações lexicais da Língua de Sinais Brasileira: uma proposta lexicográfica**. 2009, 290f. Tese (Doutorado em Linguística) – Instituto de Letras, Departamento de Linguística, Português e Línguas Clássicas, Universidade de Brasília, Brasília, 2009.

FERREIRA-BRITO, Lucinda. **Por uma gramática de língua de sinais**. Rio de Janeiro: Tempo Brasileiro, UFRJ, 1995.

FERREIRA, Daniela Bethânia dos Santos. **Estudo toponímico do centro comercial de Feira de Santana-BA**: línguas orais e Libras. 2019, 186f. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Estadual de Feira de Santana, Feira de Santana, 2019.

FRANCISQUINI, Ignez de Abreu. **O nome e o lugar**: uma proposta de estudos toponímicos da microrregião de Paranavaí. 1998, 147f. Dissertação (Mestrado em Letras) – Centro de Letras e Ciências Humanas, Universidade Estadual de Londrina, Londrina, 1998.

JESUS, Carlos Messias. **Estudo toponímico dos bairros de Feira de Santana-BA**: línguas orais e Libras. 2019, 169f. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Estadual de Feira de Santana, Feira de Santana, 2019.

LILLO-MARTIN, Daiane; GAJEWSKAY, Jon. One grammar or two? Sign Languages and the Nature of Human Language. **WIREs CognSci**, v. 5, n. 1, p. 387-401, 2014.

MARCELINO, Lico Bezerra. **Variação fonológica, morfológica e lexical em topônimos referentes a cidades acreanas**. Trabalho de Conclusão de Curso (Graduação em Letras Libras) – Universidade Federal do Paraná, Curitiba, 2021.

MELLO, Octaviano. **Topônimos amazonenses** – nomes das cidades amazonenses, sua origem e significado. Manaus: Edições Governo do Estado do Amazonas, 1967.

MIRANDA, Roselba Gomes de. **Toponímia em Libras**: descrição e análise dos sinais dos municípios de Tocantins. 2020, 186f. Dissertação (Mestrado em Letras) – Universidade Federal de Tocantins, Porto Nacional, 2020.

OLIVEIRA, Agenor Lopes de. **Toponímia carioca**. Rio de Janeiro [Distrito Federal]: Secretaria de Educação e Cultura, 1957.

PERNISS, Pamela. **Space and iconicity in German Sign Language (DGS)**. 2007. Disponível em: <http://hdl.handle.net/2066/30937>. Acesso em: 5 maio 2020.

POLGUÈRE, Alain. **Lexicologia e semântica lexical**. Noções fundamentais. São Paulo: Editora Contexto, 2018.

QUADROS, Ronice Müller de. **Libras**. São Paulo: Parábola Editorial, 2019.

SAMPAIO, Theodoro. **O tupi na geografia nacional**. Salvador: Câmara Municipal, 1901.

SOUSA, Alexandre Melo de. Metodologia para a pesquisa toponímica em língua Brasileira de Sinais. *In*: SOUSA, Alexandre Melo de; GARCIA, Rosane; SANTOS, Tatiane Castro dos. **Perspectivas para o ensino de línguas 2**. Rio Branco: Neplan, 2018. p. 9-37.

SOUSA, Alexandre Melo de. **Toponímia em Libras**: pesquisa, ensino e interdisciplinaridade. São Paulo: Pimenta cultural, 2022.

SOUSA, Alexandre Melo de. Toponímia em Libras dos bairros de Rio Branco: análise da estrutura dos sinais toponímicos e dos aspectos motivacionais. *In*: ISQUERDO, Aparecida Negri (Org.) **Toponímia urbana no Brasil**. Estudos. v. 3. Campo Grande: Editora UFMS, 2023. p. 16-53. Série Toponímia. Disponível em: <https://repositorio.ufms.br/handle/123456789/5662>. Acesso em: 15 abr. 2023.

SOUSA, Alexandre Melo de. **Toponímia em Libras**. Relatório (Pós-Doutorado em Linguística Aplicada/Libras) – Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, Florianópolis, 2019.

SOUSA, Alexandre Melo de; DARGEL, Ana Paula Tribesse Patrício. Onomástica: interdisciplinaridade e interfaces. **Revista GTLex**, v. 3, n. 1, p. 7-22, 2017. Disponível em: <https://doi.org/10.14393/Lex5-v3n1a2017-1>. Acesso em: 14 abr. 2021.

SOUSA, Alexandre Melo de; DARGEL, Ana Paula Tribesse Patrício. Caminhos da Toponímia no Brasil e as contribuições de Maria Vicentina de Paula do Amaral Dick. **Revista GTLex**, v. 6, n. 1, p. 6-21, 2020. Disponível em: <https://doi.org/10.14393/Lex11-v6n1a2020-1>. Acesso em: 15 maio 2021.

SOUSA, Alexandre Melo de; QUADROS, Ronice Müller de. Toponímia em Libras: tecnologia e ensino. *In*: SIMPÓSIO IBERO-AMERICANO DE TECNOLOGIAS EDUCACIONAIS, 2019. **Anais [...]**. Araranguá: Universidade Federal de Santa Catarina, RexLab, 2019a. p. 137-146. Disponível em: <https://publicacoes.rexlab.ufsc.br/index.php/sited/article/view/131/17>. Acesso em: 10 maio 2021.

SOUSA, Alexandre Melo de; QUADROS, Ronice Müller de. O *Web Software* Toponímia em Libras: pesquisa e ensino. *In*: SOUSA, Alexandre Melo de; GARCIA, Rosane; SANTOS, Tatiane Castro dos. **Perspectivas para o ensino de línguas 3**. São Carlos: Pedro & João Editores, 2019b.

SOUSA, Alexandre Melo de; QUADROS, Ronice Müller de. Proposta de ficha lexicográfico-toponímica digital para o estudo da toponímia em línguas de sinais. **Revista Guavira**, Três Lagoas, v. 15. n. 30, p. 126-140, 2019c.

SOUZA, Kássia Mariano de. **Registro, descrição e análise motivacional dos sinais de cidades do estado de Goiás**: a Toponímia em Libras numa interface com a Linguística de Corpus. 2023. 340 f. Tese (Doutorado em Estudos Linguísticos) – Universidade Federal de Uberlândia, Uberlândia, 2023.

SOUZA-JÚNIOR, José Edinilson. Gomes. **Nomeação de lugares na língua de sinais brasileira**: uma perspectiva de toponímia por sinais. 2012. 67f. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 2012.

TAUB, Sarah. *Language from the body: iconicity and metaphor in American Sign Language*. Cambridge: Cambridge University Press, 2001.

TRASK, Robert. Lawrence. **Dicionário de linguagem e linguística**. São Paulo: Contexto, 2004.

URBANSKI, Ítalo Rullian Webster; FERREIRA, Daiane; XAVIER, André Nogueira. Contribuições aos estudos toponímicos da Libras através da análise de sinais que designam cidades brasileiras. **Revista GTLex**, v. 6, n. 1, p. 234-267, 2020.

O NOME PRÓPRIO DE PESSOA NA TOPONÍMIA URBANA: INTERFACES ENTRE PROCESSO DENOMINATIVO E IDEOLOGIA

Ana Paula Tribesse Patrício Dargel
Aparecida Negri Isquerdo

INTRODUÇÃO

Em sua função toponímica todos os nomes são próprios e a presença de nomes de pessoas na denominação de acidentes geográficos representa a intersecção entre os dois ramos da Onomástica à medida que o nome próprio, investido de função toponímica, gera o antropotopônimo, ou seja, ocorre nesse processo “o emprego do nome individual como técnica de nomeação de acidentes geográficos” (Dick, 1990, p. 285). Além disso, os antropotopônimos evocam personagens que, de alguma forma, exerceram papel significativo na localidade em particular ou na região geográfica a que elas se vinculam em geral, como primeiros proprietários da área de onde surgiu o povoado, a cidade, o bairro. Esses nomes podem fazer referência tanto a personalidades concretas quanto a figuras históricas e lendárias que povoam o imaginário popular e são convertidos em topônimos como forma de perpetuação da memória da pessoa homenageada por meio da toponímia, particularmente a urbana. Além disso, atribuir um antropotopônimo como nome de um lugar

representa uma forma de prestar homenagem a cidadãos comuns, indivíduos de expressão local, regional, nacional ou internacional.

Considerando esse cenário, este estudo tem como propósito analisar tendências evidenciadas por formações antroponímicas na toponímia urbana de cidades sul-mato-grossenses, considerando duas taxas toponímicas (Dick, 1992), a dos antropotopônimos e a dos axiotopônimos, examinando, assim, os dados toponímicos relacionados a essas duas categorias denominativas em termos quantitativos e qualitativos – eixos como projeção da personalidade homenageada e a categoria de profissões contempladas no caso dos axiotopônimos – com vistas a delinear possíveis tendências dessas categorias de topônimos na nomeação de acidentes humanos urbanos (bairros, logradouros, praças, monumentos etc.).

1. APORTES TEÓRICOS

No âmbito deste estudo considera-se, “sob a ótica dos *nomes* – formas linguísticas empregadas na identificação de lugares, das pessoas, dos objetos espaciais, de acordo com a visão de mundo ou a cosmovisão historicamente caracterizada no território [...]” (Dick, 2007, p. 465), no caso do recorte toponímico em exame, os nomes de lugares, os topônimos oficiais, que nomeiam logradouros de cidades de grande, médio e pequeno porte, sedes de municípios do estado de Mato Grosso do Sul.

Para tanto, é preciso ter em conta que o sistema de nomeação na toponímia urbana, na sua essência, reproduz a dinâmica adotada na nomeação de acidentes físicos e humanos em geral e se reporta a “um universo de sentido: a inter-relação do sistema lexical e do sistema onomástico – denominativo, como um subconjunto daquele,

ou sua forma aplicada” (Dick, 1998, p. 78). Ainda em se tratando do processo de nomeação de lugares, a mesma toponimista destaca que no jogo que se estabelece entre léxico geral e onomástica ocorre “uma duplicidade de função sintática, cuja compreensão se realiza no contexto vivenciado por seus atores” (1998, p. 78) e nesse jogo entram em cena fatores linguísticos – o léxico toponímico tem origem no léxico geral da língua, não raras vezes na parcela do léxico regional, dialetal, fonte na qual a toponímia bebe no processo de nomeação de uma nova localidade e/ou de um elemento físico (rio, córrego, montanha, cachoeira...). Assim, ainda segundo o olhar de Dick (1998, p. 78-79),

Os nomes (sistema onomástico) podem ser percebidos, assim, como a conjunção de vários fatores necessários a uma sequência expressiva: apreensão do objeto no espaço, conhecimento e percepção de seus detalhes ou constituintes, representatividade linguística do traço percebido (cognição intelectual = uso do código, e significação do elemento codificado), manifestação denominativa (aplicação de um significante ao referente específico).

Ao focalizar a questão da motivação toponímica, Trapero (1995), por exemplo, defende que no léxico toponímico a questão da motivação linguística se manifesta de forma especial em relação a outras parcelas do léxico:

[...] puede decirse que los topónimos son, en su inmensa mayoría, nombres motivados; que el significante de tal topónimo es ése e non otro porque intenta “traducir” la realidad física concreta de un lugar, caracterizada por una cualidad que tiene ja na lengua común e ese nombre se convierte entonces en topónimo (Trapero, 1995, p. 192).

Relacionando-se essa premissa à toponímia urbana, os nomes de avenidas, logradouros, praças, monumentos particularmente se configuram como nomes motivados, no caso, na grande maioria das vezes, filtrados pelo olhar do denominador, normalmente o proprietário da área loteada a quem, pela legislação municipal, é concedida a prerrogativa de sugerir os nomes dos logradouros. Nesse processo, a indicação pode ser motivada por razões, normalmente, bastante particulares – homenagem a membros da própria família, ao seu time de futebol, à religiosidade, a amigos, a suas preferências em termos de artes, à sua região de origem e, por que não, a lugares que pretende visitar, transplantando assim nomes de lugares distantes, localizados em outros países, continentes, em geral, os mais prestigiados social e economicamente.

Não pode ser desconsiderada, nesse contexto, a força exercida pelas autoridades que administram a cidade na tendência a homenagear pessoas com quem mantêm, ou vínculos afetivos, ou compromissos políticos, afeições religiosas, tendência essa que justifica a substancial presença de nomes de pessoas, de autoridades religiosas e políticas, de profissionais que de alguma forma se destacaram na história do aglomerado humano de pequeno, médio ou grande porte, em diferentes fases de sua formação.

Em estudo sobre a relação entre “nomes, topónimos e apelidos”, na toponímia galega, Boullón Agrelo (2017, p. 34) pondera que “Unha das principais causas denominativas na toponimia é o recurso aos nomes persoais, ben sexa a partir do nome do propietario ou dunha persoa que por algunha razón este a vinculada co lugar, ben a partir dun haxiónimo”. Nessa linha de raciocínio, argumenta ainda a mesma autora que “Os topónimos que proceden dos nomes de persoas denotan unha relación co territorio nomeado, normalmente de pertenza” (Boullón Agrelo, 2017, p. 35).

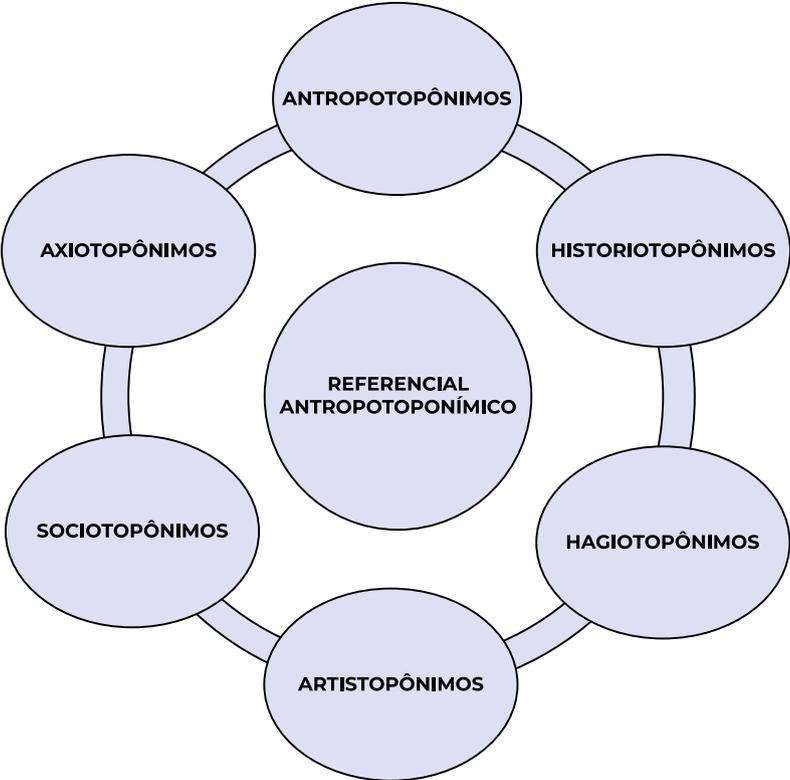
A toponímia urbana incorpora, pois, formas de percepção do espaço muito particulares, razão pela qual nomear uma rua ou uma praça com um antropotopônimo pode traduzir ideologias, autolatrias, reverências, paixões várias por pessoas investidas de poder político, de autoridades ligadas a diferentes esferas de administração, de cidadãos de grande expressão artística. Segundo sua própria dinâmica, a cidade se expande e novas motivações vão surgindo impulsionadas por fatores sociais, culturais, econômicos e, nesse processo, destaca-se o papel do denominador com a sua visão particular do espaço nomeado, suas expectativas individuais e coletivas. Desse amálgama de fatores surgem os topônimos urbanos com feições muito singulares e distintas dos designativos de elementos do meio rural.

Enfim, a natureza antroponímica dos topônimos manifesta-se, particularmente na toponímia urbana, também em graus distintos de especificidades de outras categorias de classificação toponímica (Dick, 1990), materializada pela forte presença de axiotopônimos na nomeação de logradouros em que títulos são associados a um nome de pessoa, manifestados em antropotopônimos, precedidos tanto por títulos da hierarquia militar, religiosa, quanto por títulos relativos a nomes de profissões, os sociotopônimos. Na categoria dos historiotopônimos também abundam antropônimos nomeando personagens históricos, como também entre os artistotopônimos¹ são agrupados antropônimos de vultos destacados nas artes em geral, dentre outras, a homenagem a expoentes da literatura na nomeação de ruas de determinados bairros. A Figura 1, na sequência, representa essa inter-relação que tem como núcleo a antroponímia sedimentando o referencial

¹ Taxe proposta por Isquierdo e Dargel (2020, p. 254; 268), “termo formado por um elemento base, no caso, *artis* (de origem latina) + o termo *topônimo*” (p. 254); “para classificação de topônimos com referência às artes de forma em geral (títulos de livros, nomes de artistas, poetas, escritores, esculturas, escultores, pintores e pinturas etc.)” (p. 268).

antropotopônimo (Dick, 1997) e suas consequentes relações com diferentes subcategorias toponímicas.

Figura 1 – Referencial antropotopônimo e suas relações com taxes toponímicas formadas por antropônimos



Fonte: elaboração das autoras

Não pode ser desconsiderado nesse contexto que as cidades crescem e novas motivações vão aflorando, fenômeno que permite estudos de natureza histórica comparando-se tendências toponímicas com as diferentes fases da formação das cidades. Nesse particular e a título de ilustração, menciona-se o estudo de Bittencourt (2015)

que, no estudo da toponímia urbana da cidade de Três Lagoas, mesorregião Leste de Mato Grosso do Sul, fronteira com o estado de São Paulo do qual é separada pelo rio Paraná, então com 100 anos de emancipação, analisou a relação entre a toponímia e a história social da cidade, considerando para tanto três sincronias. A autora constatou que, “Nos primeiros anos de povoamento os nomes dos logradouros eram de cunho histórico, homenagens a estados, cidades vizinhas e acontecimentos em âmbito nacional”. Numa segunda fase, quando a vila ganhou *status* de cidade, “aquelas pessoas que foram peças chaves no desenvolvimento do município passaram a ser homenageadas, demonstrando o papel da toponímia como resgate das memórias do povo”. Nessa fase a toponímia reflete aspectos da “história da rua, dos moradores, dos trabalhadores, daqueles que tiveram a rua como personagem de sua vida”. Por fim, na última sincronia estabelecida pela autora, constata-se que, ao contrário das anteriores, “a toponímia das ruas que foram nomeadas na fase do desenvolvimento desenfreado da cidade, sem representar qualquer valor histórico ou social da cidade” (Bittencourt, 2015, p. 222).

Enfim, concluiu a autora que,

[...] em sua maioria, os topônimos seguem um percurso histórico social, estão interligados a aspectos de cunho cultural, social, ideológico e linguístico de um povo; são capazes de resgatar as memórias de um grupo e refletem por meio da história como o homem viveu e vive em sociedade, como o ato de nomear revela as inúmeras facetas da realidade física e cultural de diferentes sociedades (Bittencourt, 2015, p. 223).

2. QUESTÕES METODOLÓGICAS

No âmbito do Projeto ATEMS (Atlas Toponímico do estado de Mato Grosso do Sul) já foram produzidas oito dissertações e uma tese sobre toponímia urbana, assim distribuídas: cinco sobre a toponímia urbana da cidade de Campo Grande – Oliveira (2014); Cavalcante (2019); Amorim (2017); Quisnau (2018) e Neves (2019). Além dessas, Bittencourt (2015) estudou a toponímia da cidade de Três Lagoas; Dutra (2019) da cidade de Paranaíba e Souza (2006) que analisou dados da toponímia rural e urbana de três municípios sul-mato-grossenses – Jardim, Bela Vista e Guia Lopes da Laguna. Por fim, Cazarotto (2019), na sua tese de doutorado, examinou a toponímia urbana da cidade de Ponta Porã, como parte da sua pesquisa sobre a toponímia na área de fronteira Brasil/Paraguai².

Considerando que este estudo tem como principal objetivo analisar tendências de formações antroponímicas na toponímia urbana, foram estabelecidos três eixos como fios norteadores:

1. *Produtividade* de antropotopônimos e de axiotopônimos na toponímia das cidades selecionadas (a inclusão dos axiotopônimos considerou o fato de essa taxa contemplar um antropotopônimo precedido de um título) – Quadros 1 e 2;
2. *Projeção* da pessoa homenageada pela toponímia em nível local, regional, nacional, internacional;
3. *Categoria* profissional das personalidades homenageadas.

Para tanto, como já assinalado, tomam-se amostras de dados da toponímia urbana da capital Campo Grande e de cidades do interior

² Trabalhos disponíveis no site do Projeto ATEMS: <https://atems-biblioteca.ufms.br/>.

do estado de Mato Grosso do Sul, extraídas de pesquisas sobre toponímia urbana vinculadas ao Projeto ATEMS.

3. DISCUSSÃO DOS DADOS

Na toponímia urbana, geralmente, os nomes dos logradouros não são atribuídos de forma espontânea porque ou são escolhidos pelo dono da terra do loteamento ou pelo órgão oficial. Segundo Dick (1997, p. 193), o referencial antropotoponímico geralmente se manifesta em topônimos que aludem a “indivíduos simples, o morador da rua ou do canto, o proprietário de um chão ou mesmo uma autoridade civil ou religiosa”. Nesse sentido, topônimos dessa natureza podem se reportar a nomes de autoridades, com ou sem título, artistas, políticos, artistas, arquitetos, benfeitores, religião, profissionais liberais, cidadãos comuns...). O quadro, a seguir, traz um panorama acerca da recorrência a antropotopônimos na toponímia urbana de cidades sul-mato-grossenses.

Quadro 1 – Amostra de antropotopônimos na toponímia urbana de cidades de Mato Grosso do Sul

AUTOR	CIDADE	PORCENTAGEM
SOUZA (2006)	Bela Vista	51,7%
SOUZA (2006)	Jardim	30,5%
SOUZA (2006)	Guia Lopes da Laguna	50%
SOUZA (2006)	Nioaque	17,3%
OLIVEIRA (2014)	Campo Grande (Região Urbana Centro)	41%
BITTENCOURT (2015)	Três Lagoas	45,6%
CAVALCANTE (2016)	Campo Grande (Região Urbana Imbirussu)	34,5%

AUTOR	CIDADE	PORCENTAGEM
AMORIM (2017)	Campo Grande (Região Urbana do Segredo)	47%
QUISNAU (2018)	Campo Grande (Região Urbana do Anhanduizinho)	69%
NEVES (2019)	Campo Grande (Região Urbana do Prosa)	35%
CAZAROTTO (2019)	Ponta Porã	30,1%
DUTRA (2020)	Paranaíba	71%

Fonte: elaboração das autoras

Conforme os dados do quadro 1, a taxa dos antropotopônimos é bastante produtiva na toponímia urbana de Mato Grosso do Sul. São raras as cidades do Estado com pouca incidência de antropotopônimos. Inclusive, quando são estudadas sincronias diferentes da toponímia de uma cidade, observamos a tendência de, no decorrer do tempo, topônimos não alusivos à antroponímia serem substituídos, na nomenclatura onomástica urbana, pelo nome de alguma personalidade com história de vida relevante na história da cidade. Conforme constatado neste estudo, apenas na cidade de Nioaque houve porcentagem inferior a 30% dessa categoria de topônimos, de acordo com Souza (2006).

Além dos antropotopônimos, a taxa dos axiotopônimos também sugere a intenção de registrar a natureza antroponímica na tessitura onomástica urbana. Os axiotopônimos são “topônimos relativos aos títulos e dignidades de que se fazem acompanhar os nomes próprios individuais” (Dick, 1992, p. 32). Essa categoria de topônimos pode ter como origem, em menor grau, nomeações espontâneas ou resultar de nomeações impostas, em especial as formadas com títulos relacionados à hierarquia militar e eclesiástica, muito frequentes na toponímia urbana.

Os axiotopônimos são, pois, nomes compostos formados por um título seguido de um antropônimo como ocorre em topônimos como *Coronel Camisão*, *Dona Consuelo Muller*, *Professor Candinho*, *Engenheiro Muller*, *Padre João Crippa*, identificados na toponímia urbana de Campo Grande. Nas pesquisas sobre toponímia urbana de cidades sul-mato-grossenses do interior de Mato Grosso do Sul, há nos estudos de Souza (2006) a porcentagem de 14,40% de axiotopônimos em Bela Vista; 8% em Jardim; 13% em Guia Lopes da Laguna e 15% em Nioaque³. Bittencourt (2015), por sua vez, em pesquisa na cidade de Três Lagoas, registrou 8% de axiotopônimos. Na mesma região de Bittencourt, mas na cidade de Paranaíba, Dutra (2020) catalogou 4% de axiotopônimos. Cazarotto (2019) registrou, no sul do Estado, 7,1% de axiotopônimos na cidade de Ponta Porã, fronteira com o Paraguai. Na Capital Campo Grande, Oliveira (2014) registrou 10% de axiotopônimos na Região Urbana Centro; Cavalcante (2016), 3,5% na Região Urbana Imbirussu; Amorim (2017), 4,2% na Região Urbana do Segredo; Quisnau (2018), 6% na Região Urbana do Anhanduizinho e, Neves, 5,5% na Região Urbana do Prosa.

Em se tratando de títulos no âmbito de hierarquias religiosas os axiotopônimos são formados pelo título dessa natureza. Na hierarquia eclesiástica da Igreja Católica, por exemplo, em escala descendente os três primeiros graus do sacramento da Ordem são o Episcopado (bispo), o Presbiterado (padres) e o Diaconato (diácono), logo, diácono, padres e bispos configuram-se como títulos religiosos. Os demais

³ Em relação aos dados analisados por Souza (2006), a maior produtividade de axiotopônimos em relação ao restante da amostra aqui analisada pode ser justificada pelo fato de os quatro municípios situarem-se na área geográfica que foi palco da Guerra do Paraguai (1867-1870), o que pode explicar a incidência maior de axiotopônimos formados com títulos relacionadas à hierarquia militar.

graus configuram-se como titulações eclesiásticas (dom, monsenhor, arcebispo, papa)⁴.

Nas igrejas evangélicas o *Pastor* é a autoridade máxima e a indicação varia de acordo com a congregação. Em todos os casos, o título e/ou o grau dentro da hierarquia sempre antecede o nome próprio de quem exerce a função: *Papa* Francisco, *Pastor* Moisés, *Padre* Mário. Em face disso, em termos toponímicos, consideramos as formações com *papa*, *pastor* e *padre*, por exemplo, como axiotopônimos (Quadro 2).

Quadro 2 – Amostra de *axiotopônimos* na toponímia urbana de cidades sul-mato-grossenses

AUTOR	CIDADE	PORCENTAGEM
SOUZA (2006)	Bela Vista	14,44%
SOUZA (2006)	Jardim	8%
SOUZA (2006)	Guia Lopes da Laguna	13%
SOUZA (2006)	Nioaque	15%
OLIVEIRA (2014)	Campo Grande (Região Urbana Centro)	10%
BITTENCOURT (2015)	Três Lagoas	8%
CAVALCANTE (2016)	Campo Grande (Região Urbana Imbirissu)	3,5%
AMORIM (2017)	Campo Grande (Região Urbana do Segredo)	4,2%
QUISNAU (2018)	Campo Grande (Região Urbana do Anhanduizinho)	6%
NEVES (2019)	Campo Grande (Região Urbana do Prosa)	5,5%
CAZAROTTO (2019)	Ponta Porã	7,1%
DUTRA (2020)	Paranaíba	4%

Fonte: elaboração das autoras

⁴ Fonte: <https://cffb.org.br/>. Acesso em: 25 set. 2023.

Como já pontuado neste trabalho, a toponímia urbana incorpora formas de percepção do espaço muito singulares e pode traduzir ideologias, autolatrias, reverências, paixões dos habitantes e autoridades locais por pessoas investidas de projeção política, de autoridades ligadas a diferentes esferas de poder. Os dados dos quadros 1 e 2 se enquadram no eixo um, a saber: produtividade de antropotopônimos e axiotopônimos na toponímia urbana de cidades do estado de Mato Grosso do Sul, um dos fios condutores objeto de análise deste estudo. Vale ressaltar que, ao se considerarem outras taxes como subcategorias de natureza antroponímica, é possível concluir que, na toponímia urbana de Mato Grosso do Sul, os topônimos em alusão a nomes próprios são muito recorrentes.

Nesse sentido, reitera-se que atribuir um nome de natureza antroponímica a um lugar é uma forma de prestar homenagem a pessoas que podem ter expressão local, regional, nacional ou internacional, conforme se observa nos dados do quadro a seguir com um exemplário, não só de antropotopônimos, como também de outras taxes toponímicas como historiopotônimos e axiotopônimos, (Dick, 1992) que incorporam a natureza antroponímica em sua estrutura.

Quadro 3 – Amostra de topônimos formados com nomes de pessoas *versus* projeção do personagem homenageado pela toponímia urbana de cidades de Mato Grosso do Sul

PROJEÇÃO	ELEMENTO GEOGRÁFICO	TOPÔNIMO	CIDADE	TAXIONOMIA
Internacional	Avenida	Leonardo da Vinci	Campo Grande	Artistopônimo
Internacional	Rua	Américo Vespúcio	Campo Grande	Historiotopônimo
Local	Rua	José Antônio	Campo Grande	Antropotopônimo

PROJEÇÃO	ELEMENTO GEOGRÁFICO	TOPÔNIMO	CIDADE	TAXIONOMIA
Nacional	Rua	Duque de Caxias	Bela Vista	Axiotopônimo
Regional	Praça	Ramez Tebet	Três Lagoas	Antropotopônimo
Regional	Rua	Coronel Camisão	Jardim	Axiotopônimo
Local	Rua	Egydio Thomé	Três Lagoas	Antropotopônimo
Local	Rua	Coronel Juvêncio	Guia Lopes da Laguna	Axiotopônimo

Fonte: elaboração das autoras

Conforme apontado no Quadro 3, há vultos antroponímicos de projeção internacional, nacional, regional e local como *Duque de Caxias* considerado o Patrono do Exército Brasileiro, herói da Guerra do Paraguai e amigo íntimo dos dois imperadores brasileiros. Duque de Caxias foi um nome recorrente na toponímia de Guia Lopes da Laguna, Jardim, Nioaque, Bela Vista, Campo Grande e Ponta Porã.

Outro nome de vulto histórico produtivo como topônimo no Mato Grosso do Sul é Coronel Camisão, de projeção regional por ser considerado um dos líderes militares da Guerra do Paraguai, que apareceu na tessitura onomástica das cidades de Três Lagoas, Guia Lopes da Laguna e Bela Vista. A biografia do Coronel Camisão, apesar de impressa também na história Nacional como líder e herói na Guerra da Tríplice Aliança, é mais estudada e contemplada como de projeção regional por ter enfrentado a última e mais conhecida batalha dessa guerra, a *Retirada da Laguna*, em solo sul-mato-grossense.

Nessa mesma linha, situa-se o Senador e Governador Ramez Tebet que, além de político conhecido em âmbito nacional, tem projeção e é lembrado regionalmente por ter sido político influente na cidade de Três Lagoas e no estado de Mato Grosso do Sul. Esse des-

taque regional acabou influenciando a escolha no nome da sua mãe, Angelina Tebet, para nomear uma rua da cidade de Três Lagoas.

Como topônimos de natureza antroponímica de projeção internacional, menciona-se Leonardo da Vinci, importante artista, cientista da Renascença. Além dele, aparece também na tessitura onomástica urbana sul-mato-grossense Américo Vespúcio, navegador e mercador, que a serviço do reino da Espanha desbravou o continente Americano junto com Colombo. Registra-se também o nome de Egidio Thomé como vulto de projeção local para a cidade de Três Lagoas a par de vários outros familiares, como rua Dr. Munir Thomé, por exemplo. José Garcia Leal, além de ser nome de rua de Três Lagoas e de Paranaíba, tem nomes de seus descendentes também em denominações de ruas, avenidas e escolas como em Rua Joaquim Garcia Leal (Paranaíba; Três Lagoas), Rua Protázio Garcia Leal (Três Lagoas), Escola Estadual José Garcia Leal (Paranaíba), Rua Januário Garcia Leal (Três Lagoas), Avenida Ernesto Garcia Leal (Paranaíba), Rua Manoel Garcia Leal (Paranaíba), dentre muitos outros.

Quadro 4 – Amostra de categorias de profissionais homenageadas por *antropotopônimos* na toponímia urbana de cidades de Mato Grosso do Sul

CATEGORIA	ELEMENTO GEOGRÁFICO	TOPÔNIMO	CIDADE
Artista	Rua	Anita Malfatti	Campo Grande
Artista	Rua	Belchior	Campo Grande
Professor	Avenida	Professor João Thomes	Três Lagoas
Professor	Rua	Professora Candinha	Nioaque
Arquiteto	Rua	Claudio Yoshimitsu Higa	Campo Grande
Arquiteto	Rua	Dr. Camilo Boni	Campo Grande
Engenheiro	Rua	Engenheiro Américo Carvalho Baís	Campo Grande
Engenheiro	Rua	Engenheira Anne Salim Saad	Campo Grande

CATEGORIA	ELEMENTO GEOGRÁFICO	TOPÔNIMO	CIDADE
Político	Rua	Deputado Aral Moreira	Ponta Porã
Político	Rua	Presidente Vargas	Guia Lopes da Laguna
Escritor	Rua	Clarice Lispector	Jardim
Escritor	Rua	Florbela Espanca	Campo Grande
Advogado	Avenida	Advogado Rosário Congro	Três Lagoas
Advogado	Rua	Doutor Eloy Chaves de Miranda	Três Lagoas
Médico	Rua	Dr. Munir Thomé	Três Lagoas
Médico	Rua	Dr. Jair Garcia	Campo Grande
Comerciante	Rua	Antônio Secco Thomé	Campo Grande
Comerciante	Rua	Manoel Jorge	Três Lagoas
Religioso	Rua	Padre João Crippa	Campo Grande
Religioso	Rua	João Paulo II	Ponta Porã
Militar	Rua	Almirante Barroso	Ponta Porã
Militar	Rua	General Nepomuceno Costa	Campo Grande
Vulto histórico	Rua	Tiradentes	Campo Grande
Vulto histórico	Rua	Coronel Camisão	Três Lagoas
Patrotoponímico (famílias pioneiras/ proprietários de terras)	Avenida	Eduardo Elias Zahran	Campo Grande
Patrotoponímico (famílias pioneiras, proprietários de terras)	Rua	José Garcia Leal	Paranaíba

Fonte: elaboração das autoras

Como eixo três de análise dos topônimos de natureza antropônímica, destacam-se no Quadro 4 algumas categorias profissionais recorrentes na toponímia urbana em cidades de Mato Grosso do Sul, em sua maioria em nomes de ruas, o que confirma ser “A rua [...] o ponto singular de atração da cidade, um verdadeiro microcosmo

dentro do organismo maior do aglomerado urbano. Para ela tudo converge, desde o fato corriqueiro do dia a dia, o simples entra e sai das casas até as grandes comemorações solenes e festivas” (Dick, 1997, p. 32). Enfim, as categorias perpetuadas em nomes de logradouros evocam a ideologia e preferências do denominador, sejam nomeações espontâneas ou frutos de imposição política.

Como categorias de personalidades mais produtivas na toponímia urbana de cidades sul-mato-grossenses já estudada, as pesquisas apontam para a tendência de perpetuação de nomes de militares de projeção nacional, regional ou local, como em Rua *Ricardo Franco* (patrono engenharia militar/Campo Grande), Rua *General Nepomuceno Costa* (Campo Grande), Rua *Almirante Barroso* (Ponta Porã), Rua *Coronel Camisão* (Três Lagoas, Guia Lopes da Laguna). A incidência de topônimos formados por nome/títulos de militares é bastante produtiva no estado de Mato Grosso do Sul o que se justifica pelo fato de esse território ter sido palco de uma das mais emblemáticas batalhas bélicas do Brasil, a Guerra do Paraguai (1864-1870). Dessa forma, há axiotopônimos formados com as patentes soldado, cabo, tenente, capitão, coronel, marechal, além dos títulos de conde, duque, dom, princesa, infante.

Dick (1990, p. 296-297) organiza a estrutura dos antropotopônimos com formações com: prenome: *Abel*; hipocorístico: *Bentinho*; prenome + alcunha: *José Velho*; sobrenome de família: *Barbosa*; prenome + sobrenome de família: *Antônio Borges*; onomástico completo: *Antonina de Castro Faria*. Há exemplos com todas essas formações na nomenclatura onomástica sul-mato-grossense e também formações patronímicas com topônimos formados a partir dos nomes de pioneiros como *Barboza*, *Coelho*, *Mancini*, *Thomé*, *Garcia Leal*, *Amim*, *Zahran*, *Báis*. O Quadro 4, antes apresentado, traz exemplos de categorias pro-

fissionais mais recorrentes na toponímia urbana sul-mato-grossense, contemplando, assim, o eixo 3 da proposta deste estudo.

CONSIDERAÇÕES FINAIS

A amostra de dados apresentada neste trabalho e os aspectos discutidos dão mostras da produtividade de antropotopônimos na toponímia urbana aqui focalizada. Trata-se de uma categoria toponímica que se sobressai entre as de natureza antropocultural, ainda com maior intensidade na toponímia urbana que, por incorporar formas distintas de percepção do espaço, incorpora nas formas de nomeação desse espaço tendências não observadas na toponímia rural. Assim, a significativa produtividade de antropotopônimos representa formas de homenagem a personalidades que, de alguma forma, exerceram/exercem influência na história social da cidade.

Nesse sentido, o estudo referendou a tese de que o ato de atribuir um nome de pessoas a uma rua, a uma praça pode traduzir homenagens que materializam ideologias, autolatrias, reverências, paixões várias por pessoas investidas de poder político, de autoridades ligadas a diferentes esferas de poder, de cidadãos de grande expressão artística, cidadãos comuns, enfim, são frutos de convicções individuais e sociais do denominador. A questão ideológica exerce um peso significativo no processo de nomeação de logradouros. Destacamos alguns casos: 1) perpetuação de patronímicos de várias gerações: Campo Grande: famílias Baís, Zahran, Barbosa, Coelho...; Três Lagoas e demais cidades da mesorregião Leste: famílias Tebet, Garcia Leal, Thomé...; 2) preconceito na toponímia: ínfima valorização de antropônimos de origem japonesa e paraguaia/espanhola na toponímia de Campo Grande e de Ponta Porã, se comparada à importância desses povos na formação étnica da população e na economia local; 3)

inexpressiva presença de nomes femininos/santas femininas na toponímia urbana de cidades de Mato Grosso do Sul.

Acresce-se ainda a questão da mudança de nomes de ruas por interesses políticos, como o ocorrido com ruas de Campo Grande como: Rua Velha → Rua Afonso Pena → Rua 26 de Agosto; Rua Y Juca Pirama → Marechal Rondon/Cândido Mariano; Rua da Saudade → Rua da Saudade Joaquim da Silva Sobrinho (família pioneira, latifundiário...). Há ainda a questão da interferência pessoal de autoridades na nomeação de logradouros, bairros, parcelamentos, como atestam os depoimentos do ex-prefeito de Campo Grande, latifundiário e pecuarista de grande expressão econômica, Lúdio Coelho, ex-prefeito de Campo Grande:

Bairro Zé Pereira: “Eu comprei 170 hectares de um compadre meu, o Zé Pereira, e do Metelo. Eles eram homens muito duros de negócio. Foi um negócio difícil, mas nós compramos bem barato! Aí eu falei: —Vou colocar o nome do Zé Pereiral” (Buainain, 2006, p. 313).

Sintetizando, este estudo teve como propósito fornecer um panorama da produtividade de antropônimos na toponímia urbana e, por extensão, a importância do referencial antropotoponímico também como materialização de diferentes facetas da história social da cidade por meio de homenagens tanto a pessoas comuns quanto personagens de projeção política, econômica, cultural, social. É o papel da toponímia também como reveladora de ideologias e de nuances da história e da cultura de grupos sociais.

REFERÊNCIAS

AMORIM, Bianca da Silveira de. **A toponímia urbana de Campo Grande/MS: um estudo etnolinguístico da região do Segredo**. 2017. 234 p. Dissertação (Mestrado em Estudos de Linguagens) – Universidade Federal de Mato Grosso do Sul, Campo Grande, 2017. Disponível em: <https://atems-biblioteca.ufms.br/producao/>. Acesso em: 10 out. 2021.

ATEMS – Atlas Toponímico de Mato Grosso do Sul. **Dissertações e Teses**. Universidade Federal de Mato Grosso do Sul, Campo Grande. Disponível em: <https://atems-biblioteca.ufms.br/producao/>. Acesso em: 20 out. 2021.

BITTENCOURT, Karla Porto. **Toponímia urbana da cidade de Três Lagoas/MS: interfaces entre léxico, cultura e história**. 2015. 223 p. Dissertação (Mestrado em Letras) – Universidade Federal de Mato Grosso do Sul, Três Lagoas, 2015. Disponível em: <https://atems-biblioteca.ufms.br/producao/>. Acesso em: 10 out. 2021.

BOULLÓN AGRELO, Ana Isabel. Nomes, topónimos, apelidos: camiños de ida e volta. **Revista Guavira Letras**, Três Lagoas, n. 25, p. 34-49, jul./dez. 2017. Disponível em: <http://websensors.net.br/seer/index.php/guavira/article/view/480/436>. Acesso em: 20 set.2021.

BUAINAIN, Maura Simões Corrêa Neder. **Campo Grande: memória em palavras: a cidade na visão de seus prefeitos**. Campo Grande: Instituto Municipal de Planejamento Urbano, 2006.

CFFB – **Conferência da Família Franciscana no Brasil**. Disponível em: <https://cffb.org.br/>. Acesso em 26 set.2023.

CAVALCANTE, Letícia Barbosa da Silva. **Léxico toponímico urbano na cidade de Campo Grande/MS: Região do Imbirussu**. 2016. 276 p. Dissertação (Mestrado em Estudos de Linguagens) – Universidade Federal de Mato Grosso do Sul, Campo Grande, 2016. Disponível em: <https://atems-biblioteca.ufms.br/producao/>. Acesso em: 10 out. 2021.

CAZAROTTO, Suely Aparecida. **Interfaces entre toponímia brasileira e paraguaia em área de fronteira: perspectiva etnodialetológica**. 2019. 470 p. Tese (Doutorado em Letras) – Universidade Federal de Mato Grosso do Sul, Três Lagoas, 2019. Disponível em: <https://atems-biblioteca.ufms.br/producao/>.

DICK, Maria Vicentina do Amaral. **A Motivação Toponímica e a Realidade Brasileira**. São Paulo: Edições Arquivo do Estado, 1990.

DICK, Maria Vicentina do Amaral. **Toponímia e Antroponímia no Brasil**. Coletânea de Estudos. São Paulo: Serviço de Artes Gráficas/FFLCH/USP, 1992.

DICK, Maria Vicentina do Amaral. **A dinâmica dos nomes na toponímia de São Paulo, 1954-1897**. São Paulo: Annablume, 1997.

DICK, Maria Vicentina do Amaral. O sistema onomástico: bases lexicais e terminológicas, produção e frequência. *In*: ISQUERDO, Aparecida Negri; OLIVEIRA, Ana Maria Pinto Pires (Orgs.). **As ciências do léxico**. Lexicologia, Lexicografia, Terminologia. Campo Grande: Editora UFMS, 1998. p. 77-88.

DICK, Maria Vicentina do Amaral. A terminologia nas ciências onomásticas. Estudo de caso: o Projeto ATESP (Atlas Toponímico do estado de São Paulo). *In*: ISQUERDO, Aparecida Negri; ALVES, Ieda Maria (Orgs.). **As ciências do léxico**. Lexicologia, Lexicografia, Terminologia. v. 3. Campo Grande: Editora UFMS; São Paulo: Associação Editorial Humanitas, 2007. p. 459-471.

DUTRA, Neidiani Alves da Silva. **Toponímia urbana de Paranaíba/MS**. 2020. 182 f. Dissertação (Mestrado em Letras) – Universidade Estadual de Mato Grosso do Sul, Campo Grande, 2020.

ISQUERDO, Aparecida Negri; DARGEL, Ana Paula Tribesse Patrício. A macrotoponímia dos municípios sul-mato-grossenses: mecanismos de classificação semântica. *In*: ISQUERDO, Aparecida Negri (Org.) **Toponímia**. Tendências toponímicas no estado de Mato Grosso do Sul. v. 2. Campo Grande: Editora UFMS, 2020. p. 228-272. Disponível: <https://repositorio.ufms.br/handle/123456789/3549>. Acesso em 10 out.2021.

NEVES, Janaína Domingues Verão das. **Toponímia urbana de Campo Grande/MS**: um estudo etnolinguístico dos nomes de logradouros da região urbana do Prosa. 2019, 250 p. Dissertação (Mestrado em Estudos de Linguagens) – Universidade Federal de Mato Grosso do Sul, Campo Grande, 2019. Disponível em: <https://atems-biblioteca.ufms.br/producao/>. Acesso em: 10 out. 2021.

OLIVEIRA, Leticia Alves Correa de. **Toponímia urbana da região central de Campo Grande/MS**: um olhar socioetnolinguístico. 2014. 111 p. Dissertação (Mestrado em Estudos de Linguagens) – Universidade Federal

de Mato Grosso do Sul, Campo Grande, 2014. Disponível em: <https://atems-biblioteca.ufms.br/producao/>. Acesso em: 10 out.2021.

QUISNAU, César Adilon Canhete. **A Toponímia urbana da região do Anhanduizinho de Campo Grande/MS: um estudo etnolinguístico.** 2019. 295 f. Dissertação (Mestrado em Letras) – Universidade Estadual de Mato Grosso do Sul, Campo Grande/MS, 2019.

SOUZA, Carla Regina de. **Toponímia e entrelaçamentos históricos na rota da Retirada da Laguna.** 2006, 233 p. Dissertação (Mestrado em Letras) – Universidade Federal de Mato Grosso do Sul, Três Lagoas, 2006. Disponível em: <https://atems-biblioteca.ufms.br/producao/>.

TRAPERO, Maximiliano. **Para una teoría lingüística de la toponímia:** estudios de Toponímia Canaria. Las Palmas de Gran Canaria: Universidad de Las Palmas de Gran Canaria, 1995.

O ATOBAH E AS NOMEAÇÕES GENÉRICAS NA HIDRONÍMIA BAIANA

Celina Márcia de Souza Abbade

INTRODUÇÃO

E estudar a toponímia baiana, em relação aos demais territórios nacionais, é uma grande vantagem, uma vez que a história da Bahia está intrinsecamente ligada à história de formação do povo brasileiro. O chamado processo civilizatório com as invasões europeias no século XVI ao Mundo Novo fez com que o Brasil fosse invadido a partir das terras baianas com a chegada de Pedro Álvares Cabral e sua frota em Porto Seguro, cidade que pertence ao Território de Identidade 27 – Costa do Descobrimento. A partir de então, mais de cinco séculos fizeram da Bahia um palco de invasões, migrações e colonizações, e os topônimos aqui encontrados revelam um pouco desse processo de formação e identidade nacional e, acima de tudo, baiana.

A necessidade e a curiosidade de se entender mais um pouco acerca da história de formação do povo brasileiro, a partir da vertente linguística, fez com que, em 2014, iniciássemos o Projeto ATOBAH – Atlas Toponímico da Bahia – desenvolvido na Universidade do Estado da Bahia. Tendo como objetivo levantar e estudar os nomes próprios de lugares que abrangem o território baiano, esse projeto segue a proposta teórica dos diversos estudos já desenvolvidos em outros

estados brasileiros, motivados pelo Projeto ATB – Atlas Toponímico do Brasil, idealizado pela professora Maria Vicentina do Amaral Dick na década de 1990 (Dick, 2007), cuja proposta era a de estudar os nomes de lugares do todo o Brasil, seguindo a perspectiva teórica do projeto desenvolvido por ela, o ATESP – Atlas Toponímico do Estado de São Paulo (Dick, 1996), pioneiro na toponomástica brasileira. Já o Projeto ATOBAH é uma das mais recentes propostas nessa perspectiva e faz parte do NEL – Núcleo de Estudos Lexicais, vinculado ao PPGEL – Programa de Estudos em Linguagem da Universidade do Estado da Bahia, cadastrado no CNPq.

A partir da Onomasiologia, parte dos estudos lexicais que se dedica ao estudo do *onoma* ‘nome’, a Onomástica estuda os nomes próprios, sejam eles de pessoas, Antroponímia, sejam de lugares, Toponímia. E essa relação íntima e profunda entre o nomeador e o nomeado explorada pelos estudos onomásticos, resulta inevitavelmente na relação homem-sociedade-cultura, típica dos estudos lexicais em diversas outras perspectivas. Logo, estudar os nomes próprios dos lugares de uma região ultrapassa o fazer linguístico, ampliando-se nos fazeres históricos, geográficos, étnicos, culturais e sociais de uma região e do povo que a habita.

Dessa forma, propomos aqui apresentar a frequência e as nomeações, na Bahia, de alguns termos genéricos existentes na sua hidronímia. O estudo da hidronímia baiana tem valor significativo para uma terra repleta de águas, cujo nome do próprio estado é um hidrotopônimo motivado pela segunda maior baía do mundo, a Baía de Todos os Santos, ou, se preferirem, o nome dado pelos autóctones, *Kirimurê* ‘grande mar interior’. Enquanto elemento de sobrevivência, a água é muito importante no processo de formação de uma comunidade. Os processos civilizatórios desde o início se formam ao redor

das águas, sejam as águas do Rio Nilo, sejam as do Rio São Francisco ou as da Baía de Todos os Santos.

Os nomes dos cursos d'água também são responsáveis por influenciarem na nomeação de outros acidentes humanos como cidades, fazendas, povoados etc. Segundo Isquerdo e Dargel (2013, p. 1), “há a tendência de o nome do rio inspirar o nome da cidade e não raras vezes instaura-se nesses contextos um amálgama entre rio e cidade em termos toponímicos que só informações históricas elucidam a questão da motivação do nome”.

Responsável também por motivar a nomeação dos hidrotopônimos, uma das vinte e sete taxas de classificação toponímica proposta por Dick (1990), a hidronímia baiana representa uma fonte de estudo para o ATOBAH. Observando-se que, diante da riqueza e da variação do sistema hidrográfico do Estado, a população acaba generalizando seus termos, a proposta aqui é a de apresentar a frequência e as nomações, na Bahia, dos principais termos genéricos existentes na sua hidronímia. Na própria literatura, muitas vezes não é fácil distinguir um rio de um riacho ou de um córrego, por exemplo. Visando certa uniformidade nas definições dos termos técnicos utilizados para os elementos genéricos do sintagma toponímico nas pesquisas desenvolvidas no ATOBAH, apresentamos aqui os termos que nomeiam a hidronímia baiana com suas respectivas definições.

1. O PROJETO ATOBAH

O Projeto Atlas Toponímico da Bahia (ATOB AH), iniciado em 2014, se propõe a gerar produtos significativos, os quais foram divididos em três etapas, conforme Abbade (2016, p. 580): **Organização** de um banco de dados informatizado a fim de reunir informações relativas à macro e microtoponímia encontrados nos 417 municí-

pios baianos que constituem o *corpus* do projeto; **Cartografia e digitalização** das cartas toponímicas; **Elaboração** do Atlas Toponímico, sempre tomando como base o modelo teórico proposto por Dick (1990; 1996). Atualmente, as pesquisas encontram-se na primeira etapa para a qual seguimos a seguinte proposta metodológica: levantamento dos topônimos que nomeiam os acidentes humanos e acidentes físicos documentados; coleta, análise e tabulação dos dados; organização dos topônimos em fichas lexicográfico-toponímicas; apresentação de resultados parciais.

Após a coleta, os topônimos são registrados em fichas, conforme modelo sugerido por Dick (2004) e adaptado pelo grupo, para serem analisados e classificados. Essas fichas estão sendo preenchidas por uma equipe de pesquisadores em sua maioria professores, composta de doutores, mestres, mestrandos e alunos de graduação em iniciação científica. Segue o modelo da ficha utilizada pelo ATOBAH.

Figura 1 – Modelo de ficha lexicográfico-toponímica do ATOBAH

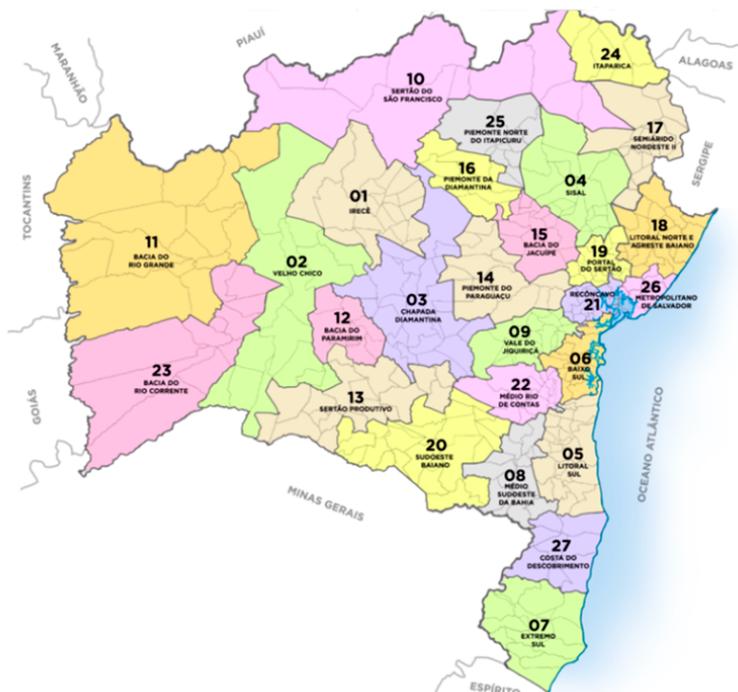
ATOBAH – Atlas Toponímico da Bahia			
TOPÔNIMO:	Corresponde ao registro do nome do lugar coletado	TAXIONOMIA:	Classificação (DICK, 1991, 1992)
LOCALIZAÇÃO:	Localização do topônimo no mapa.		
ACIDENTE:	Categoria/ tipo de acidente.		
ORIGEM:	Busca a origem do topônimo, podendo encontrar seu étimo.		
ESTRUTURA MORFOLÓGICA:	Elemento simples, composto ou híbrido.		
HISTÓRICO:	Apresenta a diacronia do topônimo, tomando por base documentos escritos.		
INFORMAÇÕES ENCICLOPÉDICAS:	Encontram-se informações várias sobre o topônimo estudado, reafirmando sua origem, estrutura morfológica e taxonomia.		
CONTEXTO:	Exemplo do uso do topônimo		
FONTE:	Apresenta-se a fonte de onde se retiraram as informações sobre o topônimo.		
COORDENADOR:			
PESQUISADOR:			
REVISOR:			
DATA DA COLETA:			

Fonte: elaborado pela autora

Nessa etapa, as fontes utilizadas para a coleta de dados são os mapas e dados encontrados na Superintendência de Estudos Econômicos e Sociais da Bahia (SEI), no Instituto Brasileiro de Geografia e Estatística (IBGE, 2017; 2018), dados do Diário Oficial do Estado (D.O.E./BA, 2008) para confirmar nomes dos recursos hídricos, uma vez que a maioria dos rios servem como limite entre uma cidade e outra. Para os acidentes humanos, acrescentamos os dados do Tribunal Regional Eleitoral da Bahia (TRE-BA, 2014), no que concerne aos nomes de povoados, distritos, assentamentos e comunidades quilombolas, uma vez que esses dados são atualizados bienalmente, diferentes dos encontrados no IBGE, cuja atualização ocorre a cada dez anos.

Os 417 municípios baianos podem ser divididos a partir de diversos critérios como: Geográfico, Turístico, Econômico, Hídrico e Territórios de Identidade. Como as pesquisas do ATOBAH são desenvolvidas com base no sentimento de pertencimento proposta pela divisão territorial desde 2010 pelo Governo do Estado da Bahia (SECULT-BA, 2010), escolhemos a divisão por Territórios de Identidade, uma vez que esse critério se propõe a considerar o que as comunidades, através de suas representações, entendem como pertencimento e identidade na formação de seus territórios. Para essa classificação, o Governo do Estado da Bahia, em 2010, identificou prioridades temáticas definidas a partir da realidade local, convidando representantes da comunidade a opinarem, o que deixa clara nossa concepção de que léxico e cultura são indissociáveis (SECULT-BA, 2010). E assim, a Bahia passou a reconhecer a existência de 27 Territórios de Identidade, constituídos a partir da especificidade de cada região e do sentimento de pertencimento de sua população. A seguir, apresentamos o mapa com essa divisão.

Figura 2 – Mapa dos Territórios de Identidade da Bahia



Fonte: Secult-BA (2010)

A relevância da hidronímia se dá, tanto nos limites entre os municípios marcados por cursos d'água, quanto na nomeação de alguns desses municípios, conforme nos mostra Correia e Abbade (2017, p. 263), ao analisarem os 27 Territórios de Identidade:

[...] dos vinte e sete, dez são nomeados a partir dos rios que os percorrem e que possibilitaram o processo de colonização e ocupação territorial, permitindo que se associe a presença da água à causa nominativa desses territórios: TI 02 - Velho Chico; TI 09 - Vale do Jiquiriçá; TI 10 - Sertão do São Francisco; TI 11 - Baquia do Rio Grande; TI 12 - Baquia do Paramirim; TI 14 - Piemonte do

Paraguaçu; TI 15 - Bacia do Jacuípe; TI 22 - Médio Rio de Contas; TI 23 - Bacia do Rio Corrente; TI 25 - Piemonte Norte do Itapicuru.

As pesquisas desenvolvidas no ATOBAH, principalmente nessa primeira etapa, constituem um trabalho em equipe para a construção do banco de dados, iniciado pelos topônimos que nomeiam acidentes físicos, mais precisamente os recursos hídricos (rios, riachos, lagoas, cachoeiras, bacias etc.). Em paralelo, os topônimos de acidentes humanos (municípios, povoados, distritos, assentamentos e comunidades quilombolas) também estão sendo catalogados. Até o momento, já foram inseridos em planilhas *Excel* os nomes de 5.349 acidentes humanos e de 9.495 acidentes físicos. Essas planilhas servem de *corpus* de base para diversas pesquisas desenvolvidas no grupo.

2. A HIDRONÍMIA BAIANA

O Estado da Bahia, com uma área territorial de 567.295 km² (IBGE, 2017), é o quarto maior do Brasil em extensão territorial, possuindo uma rede hidrográfica vasta e complexa. Para cuidar da sua hidrografia, políticas governamentais são adotadas e o planejamento da gestão das águas baianas divide o sistema hidrográfico em regiões. Essa divisão é definida pelo Conselho Estadual de Recursos Hídricos (Conerh).

Com o objetivo de orientar o planejamento e gerenciamento dos recursos hídricos em todo o país, o Conselho Nacional de Recursos Hídricos (CNRH) divide esses recursos em doze regiões hidrográficas nacionais, cuja denominação oficial é a de Regiões de Planejamento e Gestão das Águas (RPGA). Delas, duas banham a Bahia: a Região Hidrográfica Atlântico Leste que abrange quatro Estados

(Bahia, Minas Gerais, Sergipe e Espírito Santo) e a Região Hidrográfica São Francisco, sete Estados (Bahia, Minas Gerais, Pernambuco, Alagoas, Sergipe, Goiás e Distrito Federal).

A Bahia, de acordo com o Conerh, desde 2009 (Bahia, 2009), possui 26 Regiões de Planejamento e Gestão das Águas (RPGA). A partir de 2012, a configuração das regiões foi alterada passando para 25 regiões, uma vez que algumas sub-bacias foram incluídas em outra RPGA. Essa nomeação de “regiões” difere da de “bacias”, uma vez que uma região não pode ultrapassar os limites político-territoriais, se restringindo ao espaço territorial baiano, ainda que haja regiões com gestão compartilhada com outras entidades governamentais.

Assim, as RPGA, visando a uma melhor orientação no planejamento e gerenciamento das águas da Bahia, são formadas por uma bacia, uma sub-bacia, ou grupo de bacias ou sub-bacias hidrográficas vizinhas, com características naturais, sociais e econômicas semelhantes em escala regional. Outros órgãos governamentais atuam na gestão dos recursos hídricos baianos como o Instituto de Gestão das Águas e Clima (INGÁ), colegiados e os Comitês de Bacia Hidrográfica, sempre em conformidade com as diretrizes da Agência Nacional de Águas (ANA).

Outras divisões existem, como a do INGÁ que utiliza a bacia hidrográfica em sua gestão das águas. Nessa perspectiva, a Bahia em 1990 possuía 13 bacias hidrográficas que, em 1995, passaram para 10 Regiões Administrativas de Água (RAA). A primeira proposta, da década de 1990, dividia a Bahia, para fins de gestão dos recursos hídricos, em 13 regiões, chamadas de bacias hidrográficas. Com a Lei Estadual nº 6.855/95, a partir das necessidades institucionais e de revisão do sistema de gestão estadual, a Bahia foi dividida em 10 Regiões Administrativas de Água (RAA).

O Brasil é um país cuja riqueza do sistema hidrográfico é inegável. Na Bahia não poderia ser diferente. Diversos cursos d'água se fazem presentes nas terras baianas. Todas essas características refletem no processo de nomeação dos topônimos ali existentes. E, considerando a riqueza e a variação do sistema hidrográfico do Estado da Bahia, a população acaba generalizando seus termos, muitas vezes não sabendo distinguir um rio de um riacho ou de um córrego.

Entendendo **cursos d'água** enquanto corpos hídricos que deságuam em outro corpo hídrico como um mar, lago ou outro rio, podemos afirmar que existem diversas denominações para os diferentes cursos d'água. No Brasil, o que chamamos comumente de rio pode receber nomes distintos de acordo com a região do país, o fluxo de água e o local onde deságua. Entretanto, essas diversas denominações não obedecem a um critério científico ou oficial.

Assim, a **hidronímia** compreende os nomes geográficos das águas correntes ou estáveis da Terra, bem como elementos naturais ou artificiais, expostos ou submersos, contidos nesse ambiente, como arquipélago, arroio, atol, bacia, córrego, Igarapé, cachoeira, lagoa, rio, riacho, dentre outros, que nomeiam tipos distintos de acidentes hídricos.

3. O SINTAGMA TOPONÍMICO

Dentro da Onomástica, à água se dedicou um campo específico denominado hidronímia que, segundo Isquierdo e Seabra (2010, p. 88), no campo da Onomástica, é “o conjunto de acidentes geográficos que designam correntes hídricas”.

Para estudar essas nomações, encontramos o nome geográfico, ou seja, o topônimo georreferenciado, inserido em um contexto

temporal, a partir do qual podemos traçar aspectos das origens culturais e/ou históricas do elemento que ele nomeia e/ou da comunidade que o instituiu. Esse sintagma toponímico é composto pelo **termo genérico** (denominação genérica da feição geográfica que ocorre no contexto de um território, passível de ser referenciada geograficamente) e o **termo específico** (denominação específica que singulariza a identidade da feição geográfica) (IBGE, 2018). Assim, teremos o nome geográfico, termo técnico cunhado por Dick (1990), conforme figura a seguir:

Figura 3 – Nome Geográfico



Fonte: elaborado pela autora

No “Glossário dos termos genéricos dos nomes geográficos, utilizados no mapeamento sistemático do Brasil”, o IBGE (2018) registra 187 termos genéricos da base cartográfica contínua do Brasil ao milionésimo (1: 1000 000). Desses, pelo menos 64 são termos para os cursos d’água ou termos que estão relacionados às águas de alguma maneira: açude, arquipélago, arroio, atol, baía, baixa, baixão, barra, barragem, boca, boqueirão, braço, brejo, cabeceira, cabo, cachoeira, canal, corredeira, córrego, desaguadouro, enseada, grotta, grotão, hidrovia, igarapé, ilha, ilhas, ilhota, impuca, ipuca, ipueira, ipuera, lagamar, lago, lagoa, laguna, mar, oceano, olho-d’água, pontal, praia, recife, ressaca, restinga, riachão, riacho, ribeira, ribeirão, ribeirãozinho, rio, saco, salto, sanga, sangradouro, vão, vargem, varjão, várzea, vazante,

vereda, veredão, veredinha, vertente, via navegável. Esses termos estão no volume 2 da obra, pois o volume 1 foi publicado em 2015 com 140 termos. A esse novo volume, somaram-se 47 termos, totalizando os 187. Muitos deles aparecem em regiões distintas do país. Para a Bahia, 19 termos genéricos que designam os cursos d'água foram encontrados: baía, baixa, baixão, barra, barragem, boqueirão, cabeceira, cachoeira, canal, córrego, enseada, grotta, lago, lagoa, riachão, riacho, ribeirão, rio, várzea, vereda. Somam-se, a esses termos, mais três que foram encontrados nos mapas e documentos analisados, mas não aparecem no glossário do IBGE (2018): bacia, fonte e lagoinha.

Distinguir esses acidentes, nem sempre é tarefa fácil devido aos tênues limites entre alguns deles. Se pensarmos nos rios, por exemplo, muitas vezes fica difícil distingui-lo de um riacho, ribeirão ou córrego. E sabemos também que a velocidade das correntes de água, pode diferenciar, por exemplo, lagoas e lagos dos rios, córregos e outros cursos d'água, o que torna essa distinção bastante relativa.

Como um dos objetivos do ATOBAH é de estudar os termos que designam os nomes geográficos do nosso *corpus*, elencamos aqui as nomeações para esses termos genéricos encontrados até então.

Apresentamos, aqui, os cursos d'água encontrados no território baiano, todos voltados para o Oceano Atlântico, e, nacionalmente classificados em duas das doze regiões hidrográficas brasileiras: Região hidrográfica do São Francisco e Região hidrográfica do Atlântico Leste. Até então, 9.495 cursos d'água já foram catalogados pelo ATOBAH. Apenas as cachoeiras ainda não foram quantificadas, pois estamos com dificuldades de encontrá-las em documentos, uma vez que não aparecem nos mapas pesquisados. Segue a tabela com os acidentes referentes aos cursos d'água já catalogados:

Tabela 1 – Cursos d'água na Bahia

Termo Genérico	Ocorrência na BA
Riachão/ Riacho	3757
Córrego	1720
Rio	1601
Lagoa	1479
Vereda	295
Ribeirão	284
Lagoinha	139
Lago	127
Fonte	42
Barra	19
Bacia	11
Cabeceira	11
Baía	8
Boqueirão	3
Enseada	3
Total de ocorrências:	9.495

Fonte: elaborada pela autora

Para organização dos verbetes, utilizamos a seguinte estrutura: nomeação inserida em ordem alfabética, negrito, seguida da etimologia, quando possível, e da sua definição. Quando necessário, o termo foi acrescido de mais alguma informação para elucidar a sua inserção enquanto termo que designa algum tipo de acidente hídrico. As fontes consultadas para chegarmos às definições foram, principalmente: Cunha (2010), Houaiss e Villar (2001), IBGE (2018), Souza (2004 [1939]). A seguir, elencamos as denominações encontradas.

4. OS TERMOS QUE NOMEIAM ACIDENTES HÍDRICOS ENCONTRADOS NA BAHIA

Bacia – Do lat. tardio *baccea*. Conjunto de terras delimitadas pelos divisores de água e drenadas por um rio principal, seus afluentes e subafluentes.

Baía – Possivelmente do fr. *baie*, deriv. do cast. *bahia* e, este, do baixo lat. *baia*. Massa de água do mar que penetra em uma reentrância da costa, geralmente com estreitamento em sua entrada.

Barra – Deriv. de barro, e este de origem pré-romana ‘espécie de argila’. Entrada de um porto, foz de um rio ou riacho. Em algumas regiões pode ser um geomorfotopônimo, pois designa também bancos ou coroas de areia.

Boqueirão – Deriv. de boca, do lat. *buccam*. Termo regional utilizado na Região Nordeste para as aberturas ou gargantas estreitas, cortadas, por vezes, em serras, por onde passa um rio. Pode ser também a embocadura de um rio.

Cabeceira – Derivado de cabeça. Do lat. vulgar *capitia*, lat. class. *caput* ‘cabeça’. Nascentes de rio ou riacho. O mesmo que fonte, minadouro. Ponto onde surge o olho-d’água, que dá origem a um curso fluvial.

Cachoeira – Deriv. de cachão ‘borbotão’, do lat. *coctio* -onis ‘cozedura, fervura, borbulha’. Queda-d’água no curso de um rio, ocasionada pela existência de um degrau no perfil longitudinal da mesma.

Canal – Do lat. *canalis* ‘escavação, sulco, rego’. Local por onde escoam as águas fluviais ou marítimas. O canal apresenta-se em diferen-

tes formas na superfície terrestre, podendo ser natural ou construído pelo homem.

Córrego – Do lat. **currugus* ‘riacho’. Curso d’água corrente de pequeno porte. Ocorre em todas as regiões fisiográficas brasileiras, na maioria das Unidades da Federação.

Enseada – Deriv. de seio. Do lat. *sinus* ‘curvatura, volta’. Reentrância da costa bem aberta em direção ao mar, porém com pequena penetração deste, ou, em outras palavras, uma baía na qual aparecem dois promontórios (elevações) distanciados um do outro. Termo utilizado também para denominação de massa de água.

Fonte – Do lat. *fons, fontis* ‘nascente’. Local onde brota água do solo; nascente.

Estuário – Do lat. *aestuarium*. Desembocadura ou foz de rio alargada e extensa, como uma baía fechada e estreita, onde se misturam água doce do rio e salgada do mar ao sabor da correnteza fluvial e dos fluxos de marés do mar, normalmente com vários rios desaguando nele. Braço de mar que se forma pela desembocadura de um rio. Área ao longo da costa onde um rio se junta ao mar. Trecho de rio ou de mar que adentra na terra.

Lago – Do lat. *lacus* ‘lago’. Porção de água circundada por terras. Depressão do solo produzida por causas diversas e cheia de águas confinadas, mais ou menos tranquilas, pois dependem da área ocupada. As formas, as profundidades e as extensões dos lagos são muito variáveis. Geralmente, os lagos são alimentados por um ou mais rios afluentes. Possuem também rios emissários, o que evita o seu transbordamento.

Lagoa – Do lat. *lacuna*, de *lacus*. Depressão de formas variadas, principalmente tendendo a circulares, de profundidades pequenas e cheias de água doce ou salgada. A lagoa pode ser definida como lago de pequena extensão e profundidade, sendo permanente ou temporária.

Lagoinha – Dim. de lagoa. Lagoa pequena.

Riachão – Der. de riacho. Termo utilizado na Região Nordeste como curso d'água aumentativo de Riacho, sem alteração conceitual significativa.

Riacho – Do cast. *riacho*. Termo de ocorrência na Região Nordeste do Brasil e que se traduz num curso d'água ou corrente de água, que flui ou desemboca no oceano, num lago ou noutra curso d'água.

Ribeirão – Der. de riba 'ribeira, margem'. Riacho de maior curso e volume. Curso d'água ou corrente de água, que flui ou desemboca no oceano, num lago ou noutra curso de água. Apesar de o IBGE (2018) dizer que esse termo só ocorre nos Estados do Rio de Janeiro, São Paulo (interior), Goiás e Mato Grosso, foram encontrados 284 topônimos na Bahia com esse termo genérico.

Rio – Do lat. *rivus* -i. Curso d'água natural de grande volume de água que deságua em outro curso d'água. Corrente líquida resultante da concentração do lençol de água num vale.

CONSIDERAÇÕES FINAIS

A história da Bahia ultrapassa a história estadual, permeando a nacional, afinal, o Brasil começou o seu processo de colonização pelas terras baianas e Salvador foi a primeira capital do país. Através dos

topônimos, a relação entre léxico, cultura e sociedade nos remete a traços identitários. Nomear é um ato de poder e os nomes que foram dados aos lugares que conhecemos e tomamos como nossos, diz muito da nossa história. Dessa forma, o signo toponímico se torna um elemento identitário capaz de resgatar a memória e identidade de um povo, fortalecendo seu vínculo sociocultural. Os nomes nos dizem muito, sejam eles genéricos, sejam específicos. Neste trabalho, apresentamos apenas os termos genéricos que estão presentes na nomeação de elementos geográficos encontrados até então da hidronímia baiana.

Sabendo que os topônimos constituem marcas de identidade de um grupo e por isso integram o patrimônio cultural de um povo, os aspectos revelados a partir de um estudo onomástico/toponímico normalmente ultrapassam o fazer linguístico, permeando caminhos históricos e socioculturais dos nomeadores, resgatando e preservando a memória e a identidade de grupos que habitam e/ou habitaram o espaço em estudo. É essa a proposta do ATOBAH, seguindo a linha de pesquisa dos demais projetos que estudam os topônimos, espalhados pelos quatro cantos do país, iniciados pela professora Maria Vicentina do Amaral Dick na Universidade de São Paulo entre as décadas de 1980 e 1990. E, como afirma Dick (2007, p. 144): “A Onomástica, porém, é muito mais do que um mero fator auxiliar do agir e do viver individual ou coletivo; é indício de rumos tomados pelos falares ao longo dos períodos históricos, de comportamentos presentes no cotidiano e de atitudes morais ou operosas valorizadas pela população”.

A proposta deste trabalho foi mais a de realizar um levantamento dos termos genéricos dos recursos hídricos encontrados na Bahia, do que fazer estudos toponímicos propriamente ditos. Mas a proposta maior foi a de trazer à tona a problemática que envolve o

sintagma toponímico, no que concerne aos termos geográficos que antecedem os específicos.

Mesmo sabendo da escassez de estudos dessa natureza no Brasil, os projetos vigentes estimulam o Projeto ATOBAH a seguir em frente e cada dia mais perseverar nesses estudos em busca do resgate de nossa memória e história a partir de nossos topônimos, indo além das contribuições linguísticas, recuperando aspectos históricos, geográficos, culturais e sociais de nosso povo, marcados ao longo do tempo pela sua onomástica.

REFERÊNCIAS

ABBADE, C. M. de S. ATOBAH - Proposta de elaboração do Atlas Toponímico da Bahia. **Caletrosκόpio**, v. 4, p. 576-588, 2016.

BAHIA (Estado); CONERH - Conselho Estadual de Recursos Hídricos. Resolução nº 43 de 02 de março de 2009. Institui a Divisão Hidrográfica Estadual em Regiões de Planejamento e Gestão das Águas. **Diário Oficial da Bahia**, Salvador, Ano 93, n. 19.928, 3 mar. 2009. Disponível em: <http://www.seia.ba.gov.br/legislacao-ambiental/resolucoes/resolu-o-conerh-n-43>.

BRASIL, Ministério do Desenvolvimento Regional. **Agência Nacional de Águas e Saneamento Básico (ANA)**. Disponível em: <https://www.gov.br/ana/pt-br/assuntos/gestao-das-aguas/panorama-das-aguas/regioes-hidrograficas>. Acesso em: 27 abr. 2020.

CORREIA, C. M. P.; ABBADE, C. M. de S. Presença abençoada ou ausência sentida: a água na toponímia da Bahia. **A Cor das Letras**, Feira de Santana, v. 18, n. 2, p. 260-270, maio/agosto 2017. Disponível em: <http://periodicos.uefs.br/index.php/acordasletras/index>. Acesso em: 14 abr. 2021.

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

DICK, M. V. de P. do A. **A motivação toponímica e a realidade brasileira**. São Paulo: Edições Arquivo do Estado, 1990.

DICK, M. V. de P. do A. Atlas toponímico: um estudo de caso. **Acta Semiótica e Linguística**, São Paulo, v. 6, p. 27-44, 1996.

DICK, M. V. de P. do A. Rede de conhecimento e campo lexical: hidrônimos e hidrotopônimos na onomástica brasileira. *In*: ISQUERDO, A. N.; KRIEGER, M. da G. **As ciências do léxico**. Lexicologia, Lexicografia, Terminologia. v. 3. Campo Grande: Editora da UFMS; São Paulo: Humanitas, 2004. p. 121-130.

DICK, M. V. de P. do A. Atlas Toponímico do Brasil: Teoria e Prática II. **Revista Trama**, Cascavel, v. 5, p. 144-155, 2007. Versão eletrônica. Disponível em: www.unioeste.br/saber. Acesso em: 16 set. 2020.

HOUAISS, A.; VILLAR, M. de S. FRANCO, F. M. de M. **Dicionário Houaiss da Língua Portuguesa**. Rio de Janeiro: Instituto Antônio Houaiss de Lexicografia, 2001.

IBGE - Instituto Brasileiro de Geografia e Estatística. **Glossário de termos genéricos dos nomes geográficos utilizados no mapeamento do Brasil**: coordenação de cartografia. 2. ed. Rio de Janeiro: IBGE, 2018.

IBGE - Instituto Brasileiro de Geografia e Estatística. **Território e Ambiente-Bahia- Brasil**. v. 4.6.10, 2017. Disponível em: <https://cidades.ibge.gov.br/brasil/ba/panorama>. Acesso em: 21 jun. 2021.

ISQUERDO, A. N.; DARGEL, A. P. T. P. Hidronímia e toponímia: interinfluências entre meio ambiente e história. *In*: ENCONTRO INTERMEDIÁRIO DO GT DE LEXICOLOGIA, LEXICOGRAFIA E TERMINOLOGIA DA ANPOLL - ENGTLEX, 9., 2013, Caxias do Sul - RS. **Caderno de Resumos** [...]. Caxias do Sul: Site do ENGTLEX, 2013. v. 1. p. 1- 4.

ISQUERDO, A. N.; SEABRA, M. C. T. C. de Apontamentos sobre hidronímia e hidrotoponímia na fronteira entre os estados de Mato Grosso do Sul e Minas Gerais. *In*: ISQUERDO, A. N.; BARROS, L. de A. (Orgs.). **As ciências do léxico**. Lexicologia, Lexicografia, Terminologia. v. 5. Campo Grande: Editora UFMS, 2010. p. 79-99.

SECULTBA - Secretaria de Cultura - Governo do Estado da Bahia. **Divisão Territorial da Bahia**, 2010. Disponível em: <http://www.cultura.ba.gov.br/modules/conteudo/conteudo.php?conteudo=314>. Acesso em: 20 set. 2021.

SOUZA, B. J. **Dicionário da Terra e da Gente do Brasil**: Onomástica geral da Geografia Brasileira. Coleção Reconquista do Brasil, v. 234 (2ª série). Belo Horizonte: Itatiaia, 2004 [1939].

TRE - TRIBUNAL SUPERIOR ELEITORAL. **Seções Eleitorais da Bahia** - Eleições 2014. Tribunal Regional Eleitoral Bahia, 2014. Disponível em: https://www.justicaeleitoral.jus.br/++theme++justica_eleitoral/pdfs/web/viewer.html?file=https://www.justicaeleitoral.jus.br/arquivos/tre-ba-endereco-das-secoes-eleitorais-de-toda-bahia-por-zona-e-local-de-votacao-em-2014/@@download/file/TRE-BA-endereco-dos-locais-de-votacao-com-secoes-toda-bahia-2014.pdf. Acesso em: 10 set. 2021.

FRASEOLOGISMOS NO CORPUS DO PROJETO ATLAS LINGUÍSTICO DO BRASIL

Marcela Moura Torres Paim

INTRODUÇÃO

Para estudar a língua e seu sistema, composto por um conjunto de unidades em contínua relação, é preciso levar em conta os processos de troca comunicativa, a expressão de crenças e os posicionamentos histórico-sociais, quando a língua é analisada, observando o uso que os sujeitos dela fazem. Assim, aspectos extralinguísticos, situacionais, surgem nas comunicações entre os indivíduos, segundo os contextos comunicativos presentes nas interações, interferindo na escolha das unidades linguísticas e nas combinações possíveis favorecedoras da intencionalidade discursiva.

A Fraseologia trata-se de um fenômeno multifacetado com especial destaque nas línguas, já que pode integrar o sistema linguístico que reflete especificidades culturais de uma dada comunidade linguística. Por isso, há diferentes possibilidades de abordagens de estudo fraseológico, levando em conta sua relevância no campo da constituição morfológica, na articulação do discurso ou na correlação que estabelece com o mundo exterior, aspectos esses que levam as pesquisas de natureza gramatical, histórica e discursiva a voltarem-se para os fraseologismos como forma de compreender essas estruturas polilexicais.

As pesquisas linguísticas, no âmbito da Fraseologia, têm aumentado nos últimos anos, adquirindo notoriedade e rigor científico, devido ao interesse, por parte dos estudiosos, em compreender de que forma essas expressões se articulam nas línguas. Como expressão representativa desse interesse, é possível citar o Projeto CAPES-COFECUB 838/15, intitulado VALEXTRA – *Variação Lexical: teorias, recursos e aplicações: do condicionamento lexical às construções pragmáticas*, com o objetivo de realizar a descrição, a análise estrutural e contextual dos fraseologismos presentes no *corpus* do Projeto Atlas Linguístico do Brasil (ALiB) que se encontram publicados, em forma de exemplário, na obra *Nas trilhas da Fraseologia a partir de dados orais de natureza geolinguística*, organizada por Paim, Sfar e Mejri (2018).

Este texto discute resultados de pesquisa sobre a Fraseologia na Dialetoлогия, com base nos dados do Projeto ALiB, oriundos das capitais brasileiras, estratificados por sexo, faixa etária e nível de escolaridade. A Fraseologia está sendo aqui concebida como o fenômeno da linguagem que se verifica por meio de associações sintagmáticas recorrentes (Mejri, 1997). Parte-se do princípio de que as unidades fraseológicas são combinações de unidades léxicas, relativamente estáveis, com certo grau de idiomaticidade, formadas por duas ou mais palavras. No que se referem aos fraseologismos, os dados de estudos realizados mostram que as estruturas revelam a polilexicalidade e uma expressão cristalizada, cujo sentido geral não é literal. Dessa forma, as designações enfocadas possibilitam o registro da diversidade lexical do português falado nas capitais brasileiras, seguindo os princípios da Geolinguística Pluridimensional.

Inicialmente, apresenta-se uma breve fundamentação teórica acerca da Fraseologia, com foco na perspectiva francesa. Na sequência, expõe-se a metodologia de como os fraseologismos foram estudados em dados dialetológicos a partir das respostas coletadas por

meio da aplicação do questionário semântico-lexical (QSL) do Projeto ALiB, revelando a diversidade de unidades fraseológicas no português brasileiro falado.

1. OS ESTUDOS FRASEOLÓGICOS NA PERSPECTIVA DA CORRENTE FRANCESA

Para contextualizar as bases teóricas utilizadas neste trabalho é preciso, primeiramente, expor um conceito acerca da Fraseologia. Assim,

[...] a Fraseologia é um fenômeno linguístico que se relaciona com todos os níveis da linguagem (desde o fonético-fonológico ao discursivo pragmático) com o objetivo de estudar as combinações de unidades léxicas estáveis e com certo grau de idiomatidade, que sejam polilexicais, ou seja, compostas por mais de um item, e que constituam a competência discursiva dos falantes (Paim; Sfar; Mejri, 2018, p. 33).

Levando em conta essa definição, é possível verificar que os fraseologismos presentes no *corpus* do Projeto ALiB mostram que sua utilização representa parte do conhecimento linguístico do falante e que eles são mais comuns do que se pode imaginar. A observância de seu uso, também, nos possibilita conhecer a realidade linguística brasileira, tanto em termos diatópicos quanto diastráticos.

Segundo a corrente francesa, adotada por Salah Mejri (1997), a Fraseologia é uma das disciplinas que estuda o léxico, adotando como principal critério, para a consideração de um elemento como unidade fraseológica, o da polilexicalidade. Mejri (1997) aborda também o estudo do processo de fixação (*figement*) dessas unidades, contemplando vários elementos e elucidando como o processo de fixação de uni-

dades sintagmáticas livres tornam-se unidades sintagmáticas que não podem ser dissociadas. Sendo assim,

Le figement est en effet important à plus d'un titre: il engage toutes les dimensions du système linguistique (phonétique, syntaxe, [morphologie, prosodie, sémantique, etc.]). Une séquence [...] couramment employée dans la conversation de tous les jours, illustre parfaitement l'imbrication de tous les niveaux que nous venons de mentionner (Mejri, 1997, p. 23).

Essa sequência tratada pelo autor é a unidade fraseológica, ou fraseologismo e, para completar essa abordagem, destacamos que a fixação não deve ser considerada isoladamente, mas também deve ser levada em conta a estabilidade dessas unidades léxicas e o grau de idiomaticidade, além de observar se elas são usadas de forma convencional em contextos específicos.

A principal característica dessas unidades é o grau de coesão, pois os elementos que compõem a expressão perdem sua significação individual e o conjunto recebe uma nova que pode ter alguma relação com as significações anteriores ou não. Mejri (1997, p. 24) argumenta que essa sequência é dita cristalizada se ela encontra uma fixidez total ou parcial de regras da combinatória sintagmática e da comutatividade paradigmática. Para o autor, a fixidez é o processo pelo qual as formações sintagmáticas têm no seu conjunto, sintaxe interna correlacionada com o significado global, ou seja, não se pode analisar uma unidade fraseológica por meio de seus itens de forma isolada, mas todos juntos como uma estrutura apenas.

Ao usar a língua, os falantes utilizam combinações de palavras entre si para expressarem seus pensamentos e interagirem com o meio em que vivem. Às vezes, essas combinações podem ser livres ou

não. Mejrí (1997) explica que passam de livres para fixas gradualmente e quase que imperceptivelmente chamando a atenção para a noção de *continuum* no tratamento delas, o que pode ser tomado como teoria para explicar como as unidades fraseológicas aqui analisadas se tornaram ou tornam parte do vocabulário dos informantes que integram o *corpus* do Projeto ALiB.

2. COMO OS FRASEOLOGISMOS FORAM ESTUDADOS EM DADOS DIALETOLÓGICOS?

A Dialetoлогия é o ramo dos estudos linguísticos que se dedica a investigar a língua com ênfase em suas variações geográficas e sociais, permitindo a documentação do patrimônio lexical e a consequente percepção de mudanças ocorridas no léxico das línguas em decorrência de transformações de caráter político, geográfico e social. Por meio do método geolinguístico, com a sistematização da distribuição geográfica de cada traço linguístico dialetal em atlas linguísticos, a Dialetoлогия possibilita o desenvolvimento de gramáticas e glossários regionais por meio da descrição de falares produzidos por investigadores da área, evidenciando a diversidade linguística, sobretudo a lexical, e constituindo material científico que atesta o caráter heterogêneo da língua.

Dialetoлогия e Geolinguística possuem o mesmo objeto – estudo de falas populares – e se distinguem pelo modo de apresentá-lo cartograficamente. A Geolinguística se dedica ao estudo de fenômenos linguísticos representados cartograficamente em atlas linguísticos, cujas atividades tiveram maior destaque ao longo do século XX e início do século XXI, estando, portanto, a serviço da Dialetoлогия enquanto método de representação dos dados em cartas linguísticas, desenvolvidas sob orientação de bases do fazer cartográfico.

A Dialetoologia Contemporânea, ou Pluridimensional, por conceber a língua como objeto de estudo enquanto instrumento que viabiliza as relações sociais entre membros de uma coletividade, leva em consideração os fatores extralinguísticos, inerentes aos falantes e com importantes implicações na fala, não se restringindo, pois, a uma análise espacial da língua. Nesse sentido, ela também leva em consideração as variáveis sociais com o intuito de identificar áreas geograficamente definidas a partir do ponto de vista dialetal.

Sobre essa questão, Cardoso (2010, p. 26) explica que:

A dialetologia, nada obstante considerar fatores sociais como elementos relevantes na coleta e tratamento dos dados, tem como base da sua descrição a localização espacial dos fatos considerados, configurando-se, dessa forma, como eminentemente diatópica.

Nesse sentido, uma Dialetoologia caracterizada pela pluridimensionalidade desenvolve-se nos estudos linguísticos impulsionados pela própria dinâmica da língua e influenciada pela dinâmica social marcada por significativas mudanças na metade do século XX. Segundo Cardoso (2002, p. 1), os aspectos sociais sempre foram considerados nos princípios metodológicos da Dialetoologia, uma vez que tomar-se o território como ponto primordial dos estudos dialetológicos não excluía a sistematização dos falantes considerando-se determinados aspectos sociais. A esse respeito, a referida autora apresenta que

A Dialectologia apresenta-se, no curso da história, como uma disciplina que assume por tarefa identificar, descrever e situar os diferentes usos em que uma língua se diversifica, conforme a sua distribuição espacial, sociocultural e cronológica. Dois aspectos fundamentais estão, pois, na sua gênese: o reconhecimento das diferenças ou das igualdades que a lín-

gua reflete e o estabelecimento das relações entre as diversas manifestações lingüísticas documentadas ou entre elas e a ausência de dados registrados, circunscritos a espaços e realidades pré-fixados (Cardoso, 2002, p. 1).

Para a estudiosa, a alteração na dinâmica demográfica e social também proporcionou a reformulação dos parâmetros geolinguísticos, o que, segundo ela, não configura uma crise da Dialectologia, e sim um momento de expansão, aquisição de novas dimensões e captação de novos pesquisadores para a área. No que se refere aos fatos lingüísticos considerados, conforme Ferreira e Cardoso (1994, p. 16), é importante mencionar que a Geolinguística é o método que melhor se adequa aos propósitos dos estudos acerca do dialeto, uma vez que se utiliza da cartografia para documentar marcas dialetais, sendo que, por meio da criação dos atlas compostos por cartas lingüísticas, são contempladas realizações fonético-fonológicas, semântico-lexicais, morfossintáticas, pragmáticas e prosódicas, além de fatores ligados a aspectos sociais relacionados a indivíduos falantes de uma mesma língua, circunscritos em um determinado território.

Nesse contexto, os fraseologismos foram estudados em dados dialetológicos extraídos do banco de dados do Projeto ALiB em diferentes áreas semânticas. Como as pesquisas desenvolvidas, no âmbito do Projeto VALEXTRA, foram de caráter qualitativo, dada a natureza do *corpus* de cunho dialetológico, buscou-se a documentação, de unidades fraseológicas, baseada em dados orais geolinguísticos.

A metodologia aplicada na pesquisa que deu origem aos dados aqui discutidos compreendeu os pressupostos estabelecidos pelas equipes de pesquisadores do Brasil (Projeto ALiB) e da França (LDI - *Lexiques, Dictionnaires, Informatique*), seguindo os critérios metodológicos da Dialectologia Pluridimensional. Assim sendo, a seleção da

amostra foi obtida de um *corpus* documentado, segundo os parâmetros metodológicos listados a seguir:

- rede de Pontos – 25 capitais brasileiras;
- questionário – QSL/121 que investiga os nomes atribuídos à *menstruação* e está assim formulada: *As mulheres perdem sangue todos os meses. Como se chama isso?* (Comitê Nacional..., 2001, p. 31). QSL/122 a qual pesquisa os nomes atribuídos a *entrar na menopausa* e está assim formulada: *Numa certa idade acaba a/o _____(cf. item 121). Quando isso acontece, se diz que a mulher _____?* (Comitê Nacional..., 2001, p. 31); QSL/131 que registra os nomes atribuídos a *filho mais moço* e está assim formulada: *Como se chama o filho que nasceu por último?* (Comitê Nacional..., 2001, p. 31);
- informantes – 200 informantes, oito em cada capital, estratificados conforme perfil de seleção do Projeto ALiB, que considerou o sexo (masculino e feminino), duas faixas etárias (faixa I – 18 a 30 anos e faixa II – 50 a 65 anos) e dois níveis de escolaridade (fundamental incompleto e universitário).

É válido destacar que a rede de pontos do Projeto ALiB é composta por 250 localidades espalhadas em todo o território nacional, contemplando todas as regiões brasileiras e todos os estados da federação. O questionário linguístico do Projeto ALiB é constituído de sete partes distintas e, neste estudo, trabalhamos com a área temática “ciclos da vida” (questões de 121 a 135) do QSL, mais especificamente, com as questões 121, 122 e 131, conforme apresentado. Verificamos, no Questionário, que todas as perguntas se fazem acompanhar da redação de como devem ser formuladas, evitando, dessa forma, possíveis distorções nas respostas obtidas, em decorrência do modo como se apurou a forma fornecida pelo informante.

Para a realização do trabalho, seguimos os passos a seguir:

- i. audição dos inquéritos específicos do *corpus* do Projeto ALiB em questão;
- ii. transcrição grafemática do recorte analisado (ciclos da vida);
- iii. consulta a dicionários;
- iv. sistematização do material coletado em planilha do Excel, permitindo, assim, a análise desse material em todas as perspectivas possíveis, com base nos dados do Projeto ALiB: diatópica, diassexual, diageracional e diastrática.

3. UMA AMOSTRA DA REVELAÇÃO DOS DADOS

Na vigência do Projeto Valeextra (2015-2018), foi produzido um exemplário, publicado na obra *Nas trilhas da Fraseologia a partir de dados orais de natureza geolinguística*, organizada por Paim, Sfar e Mejri (2018), com o intuito de elencar fraseologismos documentados nas capitais brasileiras pelo Projeto ALiB. Dessa forma, ele é destinado a estudiosos da língua portuguesa e de suas variações, bem como a interessados pelas unidades fraseológicas. No intuito de facilitar a consulta a todos os públicos, as entradas foram organizadas, em ordem alfabética, elencando as unidades fraseológicas coletadas no *corpus* do Projeto ALiB que foram fornecidas como respostas polilexicais para as questões semântico-lexicais na primeira década dos anos 2000, e pretendeu fornecer dados que poderão contribuir para a ampliação dos estudos lexicais e, também, subsidiar o ensino e a aprendizagem da língua portuguesa.

O referido exemplário foi sistematizado da seguinte forma:

1 = unidade fraseológica (apresentada conforme a ordem alfabética);

2 = informação gramatical (classificação do sintagma (nominal ou verbal) bem como a sua composição (nome + nome; verbo + nome dentre outras possibilidades de estruturas));

3 = definição (informação por meio de texto definatório acerca da unidade fraseológica em questão);

4 = localidade(s) (organizada por cidade/estado);

5 = fonte de referência (indicando o tipo do questionário, o número da questão, a área semântica e a reprodução da formulação da pergunta);

6 = contexto (exemplo, extraído do *corpus* do Projeto ALiB. Aqui, serão encontradas as abreviaturas INQ, que significa inquiridor, ou seja, o entrevistador, e INF, que diz respeito ao informante que foi entrevistado).

Com o intuito de ilustrar uma amostra do exemplário realizado, apresentam-se as unidades fraseológicas a seguir:

Amarrar o facão. Categoria gramatical: sintagma verbal (verbo + artigo + nome). Forma alternativa de se referir à fase em que a mulher já não pode ter filhos. Localidade: Belém/Pará/Região Norte do Brasil. Fonte de referência: questionário semântico-lexical/ALiB/ questão 128/área semântica: ciclos da vida: “Numa certa idade acaba a/o _____ (cf. item 121). Quando isso acontece, se diz que a mulher _____”. Contexto: INQ.: Numa certa idade acaba a menstruação. Quando isso acontece, se diz que a mulher? INF.: **Amarrou o facão** (Belém, homem, faixa etária 1, universitário) (Paim; Sfar; Mejri, 2018, p. 58).

Estar de bode. Categoria gramatical: sintagma verbal (verbo + preposição + substantivo). Forma alternativa de denominar o período em que a mulher está menstruada. Localidades: Manaus/Amazonas/Região Norte do Brasil; Belém/Pará/Região Norte do Brasil; Rio Branco/Acre/Região Norte do Brasil; Porto Velho/Rondônia/Região Norte do Brasil; Macapá/Amapá/Região Norte do Brasil; Boa Vista/Roraima/Região Norte do Brasil; Teresina/Piauí/Região Nordeste do Brasil; São Luís/Maranhão/Região Nordeste do Brasil; Fortaleza/Ceará/Região Nordeste do Brasil. Fonte de referência: questionário semântico-lexical/ALiB/questão 121/área semântica: ciclos da vida: “As mulheres perdem sangue todos os meses. Como se chama isso?”. Contexto: INQ.: As mulheres perdem sangue todos os meses. Como se chama isso? INF.: É, aqui eles chamam regra, raramente se diz menstruação ou a mulher tá naqueles dias, são essas expressões mais, ou então, eles usam um termo, assim, mais pejorativo, né, a mulher **tá de bode** (Manaus, homem, faixa etária 2, universitário); INQ.: As mulheres perdem sangue todos os meses. Como se chama isso? INF.: Menstruação. INQ.: Existem assim alguns nomes populares? INF.: **Tá de bode.** (Teresina, mulher, faixa etária 1, universitário) (Paim; Sfar; Mejri, 2018, p. 106-107).

Fim de rama. Categoria gramatical: sintagma nominal (nome + preposição + nome). Forma alternativa de se referir ao filho que nasceu por último. Localidades: Boa Vista/Roraima/Região Norte do Brasil; Rio Branco/Acre/Região Norte do Brasil; Natal/Rio Grande do Norte/Região Nordeste do Brasil. Fonte de referência: questionário semântico-lexical/ALiB/questão 131/área semântica: ciclos da vida: “Como se chama o filho que nasceu por último?” Contexto: INQ.: Como se chama o filho que nasceu por último? INF.: Derradeiro, **fim de rama.** (Boa Vista, mulher, faixa 1, universitário). INQ.: Como se chama o filho que nasceu por último? INF.: É o **fim de rama**, é o filho caçula. (Natal, homem, faixa 1, fundamental) (Paim; Sfar; Mejri, 2018, p. 127-128).

As unidades exemplificadas retiradas do *corpus* do Projeto ALiB e publicadas no exemplário aqui tomado como fonte favorecem os estudos fraseológicos visto que estão presentes em situações comunicativas contextualizadas e são compostas por elementos que perdem o significado primário e adquirem um novo sentido. Assim, o banco de dados do Projeto ALiB reúne também amostras fraseológicas que evidenciam a polilexicalidade, com tipos de estruturas distintas e com motivações diferentes, como é possível verificar, também, na cartografia dos dados fraseológicos nas capitais brasileiras:

Figura 1 – Fraseologismos para menstruação



Fonte: Banco de dados do Projeto Atlas Linguístico do Brasil

Figura 2 – Fraseologismos para menopausa



Fonte: Banco de dados do Projeto Atlas Linguístico do Brasil

Figura 3 – Fraseologismos para filho mais moço



Fonte: Banco de dados do Projeto Atlas Linguístico do Brasil

Nos dicionários consultados, foi possível verificar que Ferreira (2010) documenta *amarrar o facão*, com o sentido de “chegar ao climatério”, como brasileirismo popular da Bahia. No Houaiss (2009), encontra-se a expressão *passar o facão*, classificada como “linguagem informal”, no sentido de “dispensar alguém do emprego”, o que vem corroborar a interpretação sugerida. Ferreira (2010) e Houaiss (2009) documentam *estar de boi* e *estar de bode*, com o sentido de “menstruação”, como brasileirismo popular; registram *fim de rama*, na acepção de “o filho mais novo; o caçula; fim-de-safra; ponta-de-rama; raspade-tacho”, como brasileirismo popular, e *raspa de tacho*, com o sentido de “o filho mais novo; o caçula; fim-de-safra; ponta-de-rama”, como brasileirismo popular.

Por meio dos dados fraseológicos coletados, a partir das questões observadas no QSL/ALiB, visualiza-se a sistematização da diversidade lexical do português brasileiro, com o registro de diferentes formas linguísticas informadas pelos falantes para um mesmo referente. Isso revela que os fraseologismos estão incorporados ao léxico do português do Brasil, estão dicionarizados e se mantêm na fala popular.

Para a formulação das questões do QSL/ALiB, seguiu-se a perspectiva onomasiológica, a qual parte do conceito para a denominação, elaborada de maneira que o informante indique a variante local a partir dos semas apresentados pelo inquiridor.

Os dados de natureza geolinguística, extraídos do *corpus* do Projeto ALiB, podem trazer uma significativa contribuição para os estudos fraseológicos visto que estão atrelados a contextos comunicativos e são constituídos por elementos que perdem o significado primário e ganham um novo sentido como, por exemplo, no caso de “amarrar o facão”. O facão é um importante instrumento de trabalho

nas áreas rurais, especialmente no trabalho agrícola. Como observa Cardoso (1998, p. 112): “A simbologia do facão para o homem do campo é muito forte. Ao usá-lo preso à cintura, como geralmente o usa, o facão torna-se símbolo do trabalho, da produção, da força, da segurança e da capacidade de autodefesa”.

A procriação é vista como uma importante função da mulher, com o mesmo valor do trabalho desenvolvido no campo. Ao entrar na menopausa, ela abandona essa função, isto é, “quebra”, “amarra”, “tira da cintura”, “deixa cair” o facão, como o agricultor que não vai mais sair para o campo. Fraseologismo com o mesmo sentido, com utilização de outro elemento metafórico, é “pendurar as chuteiras”, de referência ao jogador que deixou de competir e, portanto, colocou de lado um dos objetos que utiliza na atividade. A naturalidade com que ocorre faz com que, na área, a expressão *fulana amarrou o facão* seja neutra e desprovida de qualquer conotação pejorativa, equivalendo exatamente a “entrar na menopausa”.

CONSIDERAÇÕES FINAIS

O *corpus* pesquisado possibilitou apresentar uma amostra da riqueza fraseológica de que se reveste a língua portuguesa falada nas capitais brasileiras que compõem o exemplário apresentado em Paim, Sfar e Mejri (2018).

As unidades fraseológicas, levantadas por meio dos inquéritos realizados nas capitais brasileiras, contemplam a polilexicalidade e refletem a estabilidade, como expõe Mejri (1997), de relação tão estreita entre os elementos que os leva a perderem o significado primário para adquirirem um novo sentido.

Por tudo isso, é importante registrar a continuidade dessas análises fraseológicas, dando a oportunidade de divulgar os dados do Projeto Atlas Linguístico do Brasil, aprofundando os conhecimentos da Fraseologia a partir de dados orais de natureza geolinguística.

REFERÊNCIAS

CARDOSO, Suzana Alice Marcelino. Inovação e conservadorismo no léxico rural brasileiro. *In: CONGRESSO INTERNAZIONALE DI LINGUISTICA E FILOGIA ROMANZA*, 21., 1995, Palermo. **Atti...** v. 5. Tubingen: Max Niemeyer, 1998. p. 109-119.

CARDOSO, Suzana Alice Marcelino. A geolinguística no terceiro milênio: monodimensional ou pluridimensional? **Revista do GELNE**, Fortaleza, v. 4, n. 2, p. 1-16, 2002.

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

COMITÊ NACIONAL DO PROJETO ALiB. **Atlas Lingüístico do Brasil**. Questionários 2001. Londrina: EDUEL, 2001.

FERREIRA, Aurélio Buarque de Holanda. **Novo Aurélio Século XXI**: o dicionário da língua portuguesa. 5. ed., totalmente revista e ampliada. Rio de Janeiro: Nova Fronteira, 2010.

FERREIRA, Carlota; CARDOSO, Suzana A. M. **A Dialetoлогия no Brasil**. São Paulo: Contexto, 1994.

HOUAISS, Antônio; VILLAR, Mauro de Salles. **Dicionário Houaiss da língua portuguesa**. Rio de Janeiro: Objetiva, 2009.

MEJRI, Salah. **Le figement lexical**: descriptions linguistiques et structuration sémantique. Manouba: Publications de la Faculté des Lettres de la Manouba, 1997.

PAIM, Marcela Moura Torres; SFAR, Inès; MEJRI, Salah (Orgs.). **Nas trilhas da Fraseologia a partir de dados orais de natureza geolinguística**. Salvador: Quarteto, 2018.

NOMEAÇÕES SECUNDÁRIAS EM CONTEXTOS MIGRATÓRIOS E SUA REPRESENTAÇÃO LITERÁRIA

Márcia Sipavicius Seide

INTRODUÇÃO

Neste capítulo, são analisadas nomeações secundárias em contextos migratórios lituanos ao Brasil e a países de língua inglesa e sua representação literária. Na primeira seção do capítulo, descrevem-se a fundamentação teórica e os procedimentos metodológicos da pesquisa; na segunda seção, são apresentados os resultados de pesquisa relativos às nomeações em âmbitos migratórios e, na terceira, os que concernem sua representação literária. Nesta introdução, são fornecidas informações contextuais sobre a Lituânia, a língua lituana e informações históricas sobre a primeira e a segunda onda de emigração.

Situada na região centro norte da Europa, a Lituânia é um país litorâneo banhado pelo mar Báltico cujas fronteiras fazem divisa com a Polônia, com a Letônia, com a Bielorrússia e com Kaliningrado. De acordo com dados de 2016, sua população de 2.900.787 pessoas se aproxima da população de Brasília, Distrito Federal, que contava, naquele ano, com 3.059.444 habitantes. Em termos de área, o país pode ser comparado ao estado brasileiro da Paraíba, tendo em vista que o

estado paraibano ocupa uma área de 56.585 km², pouco menos que a Lituânia cuja área se estende por 65.200 km² (Seide, 2020a).

Do ponto de vista linguístico, a língua portuguesa e a língua lituana são bem diferentes em termos de tipologia, tendo em vista que a primeira é um idioma românico que não apresenta casos e a segunda é um idioma do ramo báltico e tem casos e declinações. Os exemplos a seguir ilustram as diferenças linguísticas entre eles:

(1) Portuguese

Pedr - o

Peter – M.SG

‘Peter ‘

(2) Portuguese

O Pedr- o

ART –DEF.M.SG – Peter – M.S.G.

The Peter

‘Peter ‘

(3) Lithuanian

Petr-as

Peter –NOM.SG.

(4) Portuguese

O Pedro me vê

The Peter- SBJ me-OBJ 1 SG see – PRS.3SG.

‘Peter sees me’

(5) Lithuanian

Petr-as mane mat-o

Peter- NOM, SG. me – ACC.SG see - PRS.3SG.

‘Peter sees me ‘

(6) Portuguese

Eu vej-o o Petr-o

I- SUBJ, 1SG see –PRS.1SG, the Peter - OBJ

‘I see Peter ‘

(7) Lithuanian

Aš mat-au Petr-ą

I – NOM.1SG. see - PRS.1SG Peter-ACC.M.SG.

‘I see Peter ‘

(Petruilionė; Seide, 2018, p. 1204).

Como se pode perceber pelos exemplos, o nome próprio *Pedro*, em língua portuguesa, equivale ao nome próprio *Petras*, na língua lituana. O gênero masculino do substantivo que se expressa, em língua portuguesa, pela flexão -o, corresponde, no idioma báltico, a diferentes flexões conforme o caso, sendo considerado como primário o nominativo cuja terminação é -as. Além dessa terminação, são também utilizadas no gênero gramatical masculino -is, *lietuvis* (lituano), -ys *arklys* (cavalo), -us *lietus* (chuva) e -ius *skaičius* (número) (Ambrazas *et al.*, 1997, p. 99-100, *apud* Petruilionė; Seide, 2018, p. 1205).

Na língua lituana, as terminações prototípicas para substantivos femininos são -a e -ė no nominativo singular. Há também outros morfemas como -i, -is, -uo, mas eles não são usados em prenomes femininos (Ambrazas, 1997). Os seguintes substantivos femininos comuns e próprios servem de exemplos: *teta* (tia), *duktė* (filha), *Kristina* e *Adelė* (Petruilionė; Seide, 2020, p. 11).

Outra diferença importante concerne aos sobrenomes das mulheres no idioma lituano, nos quais diferentes morfemas são usados. Se a mulher é solteira, o sobrenome (via de regra herdado do pai) tem

uma terminação específica. Quando ela se casa, herda o sobrenome do esposo que também é flexionado de uma maneira determinada. Assim marca-se, formalmente, pela terminação do sobrenome, se a mulher é casada ou solteira: “enquanto uma filha solteira tem o sobrenome Sipavičiute, uma mulher casada com um Sipavičius, passa a chamar-se Sipavičienė” (Seide, 2020b, p. 83).

Do ponto de vista histórico e cultural, Brasil e Lituânia têm pouco em comum. Enquanto o Brasil se tornou independente de Portugal em 1822, a Lituânia se libertou do império czarista russo em 1918, porém, da década de 1940 até 1990, foi dominada pelo regime soviético. Anos depois da restauração de sua independência, em 2003, o país ingressou na União Europeia. As poucas convergências com o Brasil consistem em ambos serem majoritariamente católicos e terem sido influenciados pelos jesuítas no mesmo período. Outro ponto de convergência decorre de o Brasil ter recebido imigrantes lituanos em duas épocas: após a primeira independência lituana e depois da 2ª. Guerra Mundial (Petrulionė; Seide, 2018).

A primeira leva migratória trouxe a maioria dos lituanos que veio ao Brasil, ela começou em 1910 e se estendeu até quase 1930 tendo sido motivada, segundo Zen (2012), pela instabilidade econômica e política. Os imigrantes dessa época tinham motivos econômicos para imigrar e eram pessoas humildes oriundas do meio rural. A segunda onda, bem menos numerosa no Brasil, era formada por intelectuais, professores universitários, engenheiros e médicos, a maioria tinha nível superior completo e foi obrigada a imigrar por motivos políticos, como pessoas deslocadas de guerra. Com o ingresso da Lituânia na União Europeia, houve e há outros movimentos migratórios, contudo, eles fogem aos propósitos da pesquisa apresentada neste capítulo.

Em todas as épocas, os Estados Unidos têm sido o destino preferido dos lituanos, inclusive, no período pré-independência, os que lá viviam tiveram um papel importante de apoio ao movimento independentista. Entre os países anglófonos que os receberam se destacam o Canadá e a Austrália, este último, ao contrário do Brasil, acolheu, majoritariamente, imigrantes que haviam sido expulsos pelo governo soviético.

Independente do motivo pelo qual uma pessoa e sua família abandonam sua pátria, as dificuldades pelas quais passam são semelhantes, não obstante poderem ser agravadas de acordo com o momento histórico e econômico do país de acolhida, pela cultura dos já estabelecidos que pode ser mais aberta ou mais fechada aos estrangeiros, e também pelas barreiras linguísticas impostas pelo desconhecimento ou pouco domínio da língua oficial da nação onde passam a viver.

1. FUNDAMENTAÇÃO TEÓRICA E PROCEDIMENTOS METODOLÓGICOS DA PESQUISA

Entre os embates e dilemas que circunstanciam os contextos migratórios estão os que se referem às barreiras linguísticas e que afetam tanto os imigrantes quanto as comunidades de fala nas quais eles ingressam. Nessas situações, às vezes, línguas até então reciprocamente desconhecidas por seus falantes nativos são postas em contato. Nos casos em que existem disparidades entre os repertórios e os sistemas antropológicos em jogo, podem surgir sensações de desconforto linguístico ao longo das interações entre nativos e imigrantes. O desconforto linguístico pode ser definido como ausência de conforto e ocorre quando

não há as condições necessárias para que haja um nível satisfatório de conforto, isto acontece quando o usuário de uma língua não se sente capaz de utilizá-la para entender, interpretar e produzir enunciados nessa língua em diferentes contextos. Quando alguém utiliza uma língua estrangeira, há uma tendência a haver maior conforto na utilização de língua materna, de um lado, e, de outro, o desconforto é maior caso se trata de línguas muito diferentes entre si, como é o caso das Libras e da língua portuguesa (Albres Santiago; Andrade, 2013).

Quando um imigrante tem um nome ou sobrenome cujas características linguísticas causam estranheza no país de acolhida, pode surgir a necessidade de minimizar o desconforto linguístico causado àqueles a quem o estrangeiro precisa se dirigir em ações cotidianas. Uma alternativa é ele decidir alterar seu próprio nome para interagir com os nativos, ou, até mesmo, dar-se outro nome, completamente diferente do seu, mas que seja linguisticamente confortável aos nativos; outra é que os nativos alterem seu nome ou lhe atribua outro que eles considerem que seja mais fácil de ser usado. Em ambas as situações, de modo voluntário ou imposto, há o surgimento de uma nomeação secundária, isto é, posterior àquela pela qual o indivíduo foi primeiramente nomeado. Como se verá adiante, os produtos desses processos de nomeação secundária podem manter-se como não oficiais, ou se tornarem oficiais.

No parágrafo anterior, foram usados os adjetivos primário, secundário, oficial e não-oficial para caracterizar diferentes tipos de nomeação de pessoas. Essa caracterização se baseia nos tipos de nomeações descritas por Van Lagendonck (2007). A nomeação oficial tem valor legal, é uma obrigação dos pais e um direito do recém-nascido. Nos países cristãos, por muito tempo, foi a Igreja a única responsável pelo registro oficial dos nomes das crianças. Hoje em dia, contudo, é o Estado que tem este dever: os pais devem registrar seus filhos no

cartório de registro civil. Sejam ou não oficiais, os nomes podem ser primários ou secundários, de acordo com a ordem cronológica em que são atribuídos. Por exemplo, numa família, um bebê recém-nascido pode ganhar um apelido pelo qual é referido e, posteriormente, ganhar um nome oficial registrado em cartório. Neste caso, o apelido é não oficial e primário e o nome civil é oficial e secundário. Contudo, o mais comum é o inverso ocorrer: que alguém receba um nome oficial pelo qual é registrado e, mais tarde, lhe seja atribuído um apelido.

Os resultados de pesquisa apresentados neste capítulo se centram nas seguintes categorias de nomes próprios de pessoas: nome civil (formado por prenome e nome de família), hipocorístico, apelido e pseudônimo. Enquanto no quadro 1 (extraída de Amaral; Seide, 2000, p. 100), há definições destes diferentes antropônimos, no quadro 2 há exemplos e informação sobre se as nomeações que lhes dão origem são, de modo prototípico, primárias, secundárias, oficiais ou não oficiais.

Quadro 1 – Antropônimos analisados

Tipo	Definição e principais características
Prenome	Antropônimo que distingue o indivíduo dentro dos grupos sociais de maior intimidade. Antecede o sobrenome e pode ser simples, composto ou justaposto.
Sobrenome	Antropônimo que identifica o pertencimento do indivíduo a uma família. Geralmente provém dos genitores e sucede o prenome.
Apelido (ou alcunha, ou cognome)	Antropônimo que se atribui a um indivíduo geralmente por outra pessoa e que costuma aludir a uma característica física ou intelectual ou ainda a um fato ou comportamento social.
Hipocorístico	Antropônimo formado a partir de uma alteração morfológica (abreviação, diminutivo, aumentativo, etc.) de outro antropônimo. Geralmente criado em ambientes de maior intimidade.
Pseudônimo	Antropônimo empregado por um indivíduo em lugar do seu nome civil e escolhido pelo próprio portador do próprio nome.

Fonte: Amaral e Seide (2000, p. 100).

Quadro 2 – Tipos de nomes e de nomeações

Tipo de nome	Tipos de nomeação envolvida	exemplos
nome civil	primária e oficial	José Carlos Alvares
hipocorístico	secundária e não oficial	Zeca
pseudônimo	secundária e não oficial	Antônio Pires
apelido	secundária e não oficial	gaúcho

Fonte: Elaboração própria

Apresentados, de modo geral, os tipos de nome e os processos de nomeação normalmente envolvidos é preciso fazer considerações sobre sua representação literária, objeto de estudo da Antroponomástica ficcional. Os nomes de personagens fictícias podem ser criados a partir de seu significado etimológico ou com base na norma antroponímica em vigor no local e na época em que a obra literária está ambientada.

Quando uma personagem é nomeada em função de seu significado etimológico, cria-se uma relação entre a etimologia de seu nome e a caracterização, o comportamento e o destino da personagem no universo fictício do qual faz parte. Essa relação pode ser de conformidade ou desconformidade. No primeiro caso, sua carga semântica antecipa e simboliza a personagem. No segundo, há uma ironia por parte do criador da obra literária: a personagem comporta-se contrariamente ao que se poderia esperar tendo em vista o significado seu etimológico do seu (Amaral; Seide, 2020, p. 197).

Nos casos em que a nomeação fictícia se pauta na norma antroponímica vigente no tempo e no espaço de ambientação da obra literária, também pode haver dois tipos de relações. Há convergência quando as características do nome da personagem mimetizam os nomes que pessoas da mesma época, do mesmo lugar e que fazem

parte da mesma sociedade que a obra retrata. Há divergência quando as características do nome da personagem são diferentes da norma (Amaral; Seide, 2020, p. 199).

É preciso ressaltar que a identificação e análise das diferentes funções da antroponímia literária, numa determinada obra literária, dependem de procedimentos também distintos. Enquanto, para as funções baseadas no étimo, é necessário considerar o significado semântico e a origem dos nomes, a análise das funções pautadas em uma determinada norma antroponímica requer que ela seja conhecida de antemão para que saiba se há convergências ou divergências entre a antroponímia social e a norma antroponímia literária.

Conforme já mencionado, a pesquisa cujos resultados são descritos neste capítulo se centra na análise de nomeações secundárias e de sua representação literária em contexto de migração lituana. Enquanto a análise das nomeações secundárias foi feita a partir de relatos espontâneos de descendentes de imigrantes provenientes da Lituânia postados em grupos fechados de *Facebook* que congregam esses imigrantes e seus descendentes, a descrição de sua representação literária no romance *Circle of Amber* (Reilly, 2016), escrito por uma filha de lituanos da 2ª onda migratória, ambientado na Lituânia e na Austrália e no período que abrange do fim do século XVIII até fim do século XIX, aproximadamente. A seção a seguir apresenta e discute os resultados obtidos com dados coletados nos grupos de *Facebook*.

2. NOMEAÇÕES SECUNDÁRIAS NO CONTEXTO DE MIGRAÇÃO LITUANA

O primeiro depoimento a ser analisado é o de um neto de lituano que reside no Brasil:

Olá pessoal!! Uma ajuda, por favor, estou começando agora a juntar os documentos para solicitar a dupla cidadania [...] O problema é que mudaram o nome do meu avô de **Vincas** para **Vicente**. Sabem se pode ter algum problema? (grifos nossos).

Nesse depoimento, o informante revela que houve mudança de nome civil do imigrante lituano: no registro oficial da Lituânia seu prenome é *Vincas*, mas, no registro oficial do Brasil, mudou para *Vicente*. No idioma lituano, originalmente, *Vincas* é hipocorístico do *Vincentas* que equivale ao nome *Vicente* em língua portuguesa. Interessante notar que, na língua lituana, uma nomeação prototipicamente secundária, o apelido, ocorreu de forma primária. Posteriormente, no Brasil, houve recuperação da forma original do nome e registro em sua versão em língua portuguesa. O depoente menciona sua intenção de pedir a restituição da cidadania lituana o qual o faz prever futuros problemas ao longo deste processo: provavelmente terá que provar que, apesar das mudanças de prenome, referência é feita à mesma pessoa, isto, é ao seu avô, univocidade referencial imprescindível para a comprovação de parentesco com um cidadão lituano.

O segundo depoimento também é de um descendente brasileiro e se centra no registro, no Brasil, do sobrenome de seu bisavô. Pelo grau de parentesco e outras informações do depoimento, provavelmente, o imigrante fez parte da primeira onda de migração ao Brasil:

Tudo bem? Meu bisavô é lituano, nascido em Utena, tudo que eu tenho é a data de desembarque no porto de Santos, CIC e RNE brasileiro, também certidão de óbito e casamento. Solicitei ao “istorijos archyvas”, mas eles não encontraram o documento, eu sei que meu sobrenome BEIVIDAS é uma tradução errada e não sei o nome correto. Nem sei se o nome dele e dos pais deles estão corretos nos documentos brasileiros e nem as datas de nascimento.

De acordo com esse depoimento, na série de documentos oficiais brasileiros, consta o sobrenome Beividas que o depoente acredita estar grafado erroneamente. Se assim for há, neste depoimento, um exemplo de nomeação secundária e oficial de sobrenome. O depoente, contudo, não tem certeza se esta é a única divergência entre o registro das informações oficiais lituanas do seu bisavô: pode ser que, nos documentos brasileiros, tenha havido alteração de data e/ou de local de nascimento, tendo em vista que, pela informação prestada aos arquivos históricos lituanos, não foram encontrados documentos nos quais poderia haver informações equivalentes, como seria o caso, por exemplo, de uma certidão de batismo ou de casamento. O importante a se destacar, nesse depoimento, é a existência de nomeações secundárias oficiais de sobrenome em contexto migratório.

Diferente desses dois primeiros depoimentos, os apresentados e analisados a seguir são de descendentes de imigrantes que foram acolhidos por países de língua inglesa. O quarto depoimento é de um filho de imigrante da 2ª onda migratória cujo país de acolhida foi o Canadá:

No passado, no Canadá de 1950, nós imigrantes sentíamos uma enorme pressão para mudarmos nossos nomes. Eu nunca mudei. Meu pai sucumbiu à pressão e deixou os canadenses o chamarem de Charlie ao invés de Kazimieras / Kazys (tradução nossa).

Nesse depoimento, há informação sobre nomeação primária e oficial na Lituânia (a que resultou no nome *Kazimieras*, equivalente ao nome Casimiro, em língua portuguesa) e a do apelido, como nomeação secundária e não oficial (*Kazys*). Comparando-se esse dado com o do primeiro depoimento, fica claro que um prenome morfológicamente caracterizado como um hipocorístico, portanto, um tipo

de apelido, pode ser fruto tanto de uma nomeação primária quanto de uma nomeação secundária que pode ou não ter status oficial. Outro ponto igualmente interessante refere-se à natureza imposta da renomeação: enquanto o depoente resistiu à pressão, seu pai cedeu e mudou seu prenome para outro capaz de propiciar conforto linguístico aos canadenses que são falantes nativos de língua inglesa.

A pressão pela adoção de nomes mais adaptados à língua inglesa resultando em nomeações secundárias e oficiais é atestada por outro depoimento, no qual se informam motivos pelos quais alguns imigrantes lituanos acharam vantajoso mudar seus próprios nomes:

Também no Canadá da década de 50. Meu nome lituano era Liudgardas. A mudança para Leonard aconteceu não por causa da pressão, mas porque nenhuma das crianças nativas conseguia pronunciar o nome (tradução nossa).

Nesse depoimento, o filho de imigrante explica que os pais consideraram o fato de seu nome ser de difícil pronúncia pelas crianças com as quais ele convivia. Essa explicação indica que eles tiveram a intenção de facilitar a integração social do filho ao país de acolhida por meio da supressão do desconforto linguístico causado pelo nome lituano aos falantes de língua inglesa.

Nesse outro relato de filho de imigrante da segunda onda migratória há o seguinte testemunho de mudança oficial de sobrenome ocorrida nos Estados Unidos por vontade do portador do nome, iniciativa que o filho não consegue compreender:

Ele era Gregonis então, mudou o seu sobrenome legalmente para Gregg em 1946, ele teve que pagar \$24 pela mudança e outros \$24 dólares para mudar o nome de sua noiva, mais tarde minha mãe. Em 1946, \$48 dólares era muito dinheiro. Isto não faz

sentido para mim, com certeza eu não teria feito isto no lugar dele (tradução nossa).

Comparando esse depoimento com o primeiro analisado neste capítulo, que também relata mudança oficial de sobrenome, percebem-se diferenças. No caso do relato do descendente de lituano que mora no Brasil a mudança foi involuntária. Ao que tudo indica, foi fruto de um processo linguístico de adaptação do sobrenome ao registro em língua portuguesa por parte do responsável pelo registro dos imigrantes estrangeiros. Já, no caso do relato do descendente que vive nos Estados Unidos, houve, por parte do portador do nome, uma escolha consciente de mudança de sobrenome. Chama a atenção, também, o fato de ter havido esta possibilidade de alteração oficial de sobrenome nos Estados Unidos, enquanto, no Brasil, a criação de variantes ortográficas de sobrenome é feita involuntariamente.

No último depoimento analisado, um filho de imigrante da 2ª onda migratória que atualmente reside na Austrália faz uma narração das mudanças de nome pelas quais ele passou e fornece uma interessante descrição da sua situação atual, marcada pela convivência de diferentes nomes que são usados em contextos diferentes:

Meus pais são lituanos, mas eu cresci em países de fala inglesa (Canadá, Nova Zelândia, Austrália). Como você pode imaginar, não era muito prático tentar e insistir em ser chamado de Gintautas (a sílaba tônica é a primeira) nesses países. Então ficou "Gintis". Este é o nome que está no meu passaporte australiano, o banco lituano onde eu abri uma conta insistiu que eu usasse este nome [...] é por isso que na minha assinatura de e-mail tem Gintis na parte esquerda para falantes de inglês e Gintautas para falantes lituanos (tradução nossa).

Pelas informações dadas no depoimento, houve uma nomeação primária não oficial do prenome *Gintautas*, seguida de nomeação secundária e oficial do hipocorístico *Gintis*. Conforme observado em outros depoimentos, a nomeação secundária é motivada pelo conforto linguístico que o novo prenome pode proporcionar nos meios em que é usado. Nesse caso, parece não ter havido pressão para a mudança, mas sim um consenso. Diferente dos depoimentos anteriores, neste caso, o depoente é um homem bilíngue e bicultural que mora na Austrália, mas já viveu na Lituânia e presta serviços de tradução para clientes lituanos. Percebe-se, por sua narrativa, a importância do status oficial dos nomes. Apesar de poder usar, conforme as circunstâncias, ora seu nome primário *Gintautas*, ora seu nome secundário *Gintis* a não oficialidade de seu nome primário impediu seu registro nos dados bancários da Lituânia.

A síntese das informações coletadas nos depoimentos mostra que, no contexto de migração lituana, houve nomeação secundária oficial tanto de prenome quanto de sobrenome e que ela pode dar-se de maneira voluntária ou involuntária. Há casos em que a mudança foi feita por decisão voluntária e consciente do designador ou do portador do nome muitas vezes em decorrência de pressões sociais. Como quer que seja, em todos os casos de nomeações secundárias, elas visaram proporcionar conforto linguístico aos falantes nativos da língua oficial do país de acolhida.

Descritas as nomeações secundárias ocorridas nas sociedades que receberam imigrantes lituanos no Brasil e em países anglófonos, analisam-se, a seguir, cenas reveladoras da representação literária dessas nomeações no romance *Circle of Amber* (Reilly, 2016).

O romance de Reilly descreve como vivem três gerações de mulheres de uma mesma família de fins do século XVIII até o fim

da década de 1990. As cenas descritas a seguir são relativas à parte que narra o início da vida dos imigrantes na Austrália por volta da década de 1950. A primeira cena em que há informação sobre nomeação secundária se passa entre dois jovens lituanos que se conhecem num evento da comunidade lituana de uma grande cidade australiana. Nessa cena, o rapaz se apresenta à moça que, anos mais tarde, será sua esposa e mãe de seus filhos: “Acho que devemos nos apresentar. Meu nome é *Algimantas Kelmas*, mas os meus colegas insistem em me chamar de *Al Kelm*. Eles dizem que é mais fácil para eles. Bom, nós estamos na Austrália agora” (Reilly, 2016, p. 153, tradução nossa).

Percebe-se, na fala da personagem, uma breve narrativa de uma nomeação secundária que atingiu tanto o prenome quanto o sobrenome do imigrante lituano: seu prenome passou de *Algimantas* (nome inexistente em outros idiomas, cuja pronúncia é difícil para anglofalantes) para *Al* (um hipocorístico para nomes de língua inglesa como Albert) e seu sobrenome foi abreviado para *Kelm*, abreviação que o torna homônimo ao sobrenome de origem alemã *Kelm* (Seide, 2019). A personagem explica que é chamada assim por insistência dos colegas de trabalho, o que indica que a nomeação não foi voluntária, mas sim imposta. A justificativa dada para essa renomeação (o nome *Al Kelm* é mais fácil para eles) evidencia que o propósito era o de proporcionar conforto linguístico aos falantes nativos. A parte final dessa fala, por sua vez, revela que a atitude do imigrante perante a renomeação imposta é a de resignação: “nós estamos na Austrália agora”.

O ocorrido nessa fala vai ao encontro dos depoimentos das mudanças de nomes de *Vincas* para *Vicente*, de *Kazys/Kasimiras* para *Charlie* e de *Liudgardas* para Leonard. Vale ressaltar que, com relação aos dois últimos, também houve menção ao desconforto linguístico como a causa da renomeação. A mudança de sobrenome também foi atestada no depoimento sobre os sobrenomes *Beividas* e *Gregg*, origi-

nalmente *Gregonis*. Essas convergências evidenciam que as nomeações ficcionais dessa personagem se pautaram pela norma antroponímia e com ela converge, motivo pelo qual se pode concluir que a escolha dos nomes das personagens apresenta função mimética.

Diferente dos depoimentos, na representação literária das nomeações secundárias em contextos migratórios, são explicitados os conflitos linguístico, cultural e identitário implicados nos processos de renomeação. No caso de *Algimantas*, percebe-se o conflito e também a atitude de resignação do imigrante. Na cena, a seguir, a filha de *Algimantas Kelmas*, que se chama *Gaila*, sofre assédio moral no ambiente escolar por causa de seu nome:

Ah, você sabe exatamente como é isso né', sua burra estúpida, disse rindo *Sharon*, a valentona da turma. *Gaila* ficou vermelha como uma beterraba. Ela odiava o modo como faziam piada com o nome dela. Ela ia abreviar o nome para *Gail* o quanto antes (Reilly, 2016, p. 191, tradução nossa).

Em inglês, a palavra *galah* – cuja pronúncia é praticamente idêntica à pronúncia lituana do nome *Gaila* –, em sua acepção original, nomeia um pássaro australiano que, em língua portuguesa, se chama cacatua. A segunda acepção da palavra é exclusiva do inglês australiano. Trata-se de uma palavra pejorativa utilizada para dizer que uma pessoa não é nada inteligente. Para manter a metáfora zoonímica, traduzi “*stupid galah*” para “burra estúpida”. Considerando a pronúncia dos nomes, percebe-se que, aos ouvidos australianos, o nome *Gaila* é equiparado ao nome do pássaro socialmente avaliado como tolo, o trocadilho vem desta semelhança fonética (Seide, 2019, p. 349).

De um lado, os depoimentos sobre nomeações secundárias em contextos migratórios apontam para a existência de nomeações secundárias impostas e oficiais (adaptação linguística do nome civil em documentos no Brasil e pressão para troca de nomes em países anglófonos), impostas e não oficiais (apelidos atribuídos no ambiente de trabalho ou ambiente escolar), voluntárias e não oficiais (pseudônimos), voluntárias e oficiais (mudança oficial de nomes e sobrenomes em alguns países). De outro, no caso do romance *Circle of Amber* a representação literária das nomeações impostas aos personagens mimetiza o real e ilustra uma das funções literárias dos nomes ficcionais: a função mimética.

CONSIDERAÇÕES FINAIS

Como se evidenciou pelos dados discutidos neste capítulo, em contextos migratórios, há processos nomeadores adicionais que lhes são peculiares. O nome civil original, na maioria das vezes e, especialmente quando o imigrante é recebido num país cuja língua oficial é bem diferente da dele, é traduzido ou transliterado para a língua oficial do país de acolhida e o imigrante recebe um novo nome civil. Pode ocorrer também de o imigrante receber outro nome em seu local de trabalho e este passar a ser usado socialmente. De modo implícito ou explícito, sempre houve e continua havendo regras que determinam como podem ou devem ser registrados os nomes dos imigrantes de um país. Atualmente, há uma tendência a essas regras fazerem parte de políticas linguísticas nacionais, as quais podem ser mais restritivas ou mais liberais. Um exemplo de política linguística antroponímica liberal é a que está atualmente em vigor na França e na Hungria:

Uma mudança recente na política antroponímica da França permitiu o registro de nomes na língua húngara nos cartórios, outra mudança, na Hungria, permitiu o registro no país de pessoas nascidas no exterior. Decorrente destas mudanças, pais que emigraram por motivos econômicos, têm filhos pequenos e atitude bilíngue e bicultural optaram pelo duplo registro e pela escolha de nomes existentes em ambos os idiomas: a forma francesa do nome é registrada na França e a forma húngara na Hungria (Seide 2020c, p. 242).

Percebe-se, assim, que dependendo da política linguística de *facto* adotada em cada país, as mudanças oficiais de nomes são possíveis, impossíveis ou obrigatórias. Além disso, as possibilidades de mudança de nome em contextos migratórios indicam a necessidade de desenvolvimento de mais pesquisas que relacionem as características formais e morfológicas àquelas que são decorrentes do uso pragmático dos nomes em sociedade. Também seria interessante investigar em que medida há relação formal entre o prenome oficial secundário e o prenome primário e se há outros processos de criação de nomes secundários em contextos migratórios.

REFERÊNCIAS

ALBRES SANTIAGO, Vânia de Aquino; ANDRADE, Cristiane Esteves de. Surdez e sociedade: Questões sobre conforto linguístico e participação social. In: ALBRES, Neiva de Aquino; NEVES, Sylvia Lia Grespan. **Libras em Estudo**: Política Linguística. São Paulo: FENEIS, 2013. p. 145-163.

AMBRAZAS, Vitautas (Ed.). **Lithuanian Grammar**. Vilnius, Lithuania: Institute of the Lithuanian Language, 1997.

AMARAL, Eduardo Tadeu Roque; SEIDE, Márcia Sipavicius. **Nomes próprios**: introdução à antroponímia brasileira. São Paulo: Blucher, 2020. Disponível em: <https://openaccess.blucher.com.br/download-pdf/453>. Acesso em: 23 set. 2021.

PETRULIONĖ, Lolita; SEIDE, Márcia Sipavicius. Between languages and cultures: an exploratory study of usage of Lithuanian and Brazilian masculine anthroponyms. **Relin**, Belo Horizonte, v. 23 n. 3, p. 1201-1226, jun. 2018. Disponível em: www.periodicos.letras.ufmg.br/index.php/relin/article/view/12745. Acesso em: 23 set. 2021.

PETRULIONĖ, Lolita; SEIDE, Márcia Sipavicius. Formation and usage of hypocoristic forms in Brazilian Portuguese and Lithuanian. **Revista Alfa**, São Paulo, v. 64, p. 1-27, ago. 2020. Disponível em: <https://doi.org/10.1590/1981-5794-e11611>. Acesso em: 23 set. 2021.

REILLY, Jura. **Circle of Amber**. Higthon Victoria: Jura Reilly, 2016.

SEIDE, Márcia Sipavicius. Deslocamentos identitários e nomeação no Romance *Circle of Amber*. **Cadernos Seminal Digital**, Rio de Janeiro, v. 32, n. 32, p. 318-358, jan. 2019.

SEIDE, Márcia Sipavicius. Antroponímia, diáspora y migración: los descendientes de lituanos en Brasil. **Onomástica desde América Latina**, Cascavel, v. 1, n. 1, p. 100-121, jan. 2020a. Disponível em: <https://doi.org/10.48075/odal.v1i1.24156>. Acesso em: 23 set. 2021.

SEIDE, Márcia Sipavicius. Antroponomástica Comparada. **Onomástica desde América Latina**, Cascavel, v. 1, n. 2, p. 83-92, jul. 2020b. Disponível em: <https://doi.org/10.48075/odal.v1i2.25488>. Acesso em: 23 set. 2021.

SEIDE, Márcia Sipavicius. Do direito ao nome ao nome direito: as políticas linguísticas sob o viés antroponímico. In: BERGER, I. R; REDEL. **Políticas de gestão do multilinguismo: práticas e debates**. Campinas: Pontes, 2020c. p. 231-252.

VAN LANGENDONCK, Willian. **Theory and typology of proper names**. Berlin: Walter de Gruyter, 2007.

ZEN, Erick Reis Godilauskas. **Identidade em conflito**. Os imigrantes lituanos na Argentina, Brasil e Uruguai (1920-1950). 2012. 184 f. Tese (Doutorado em História Social) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2012.

HIDROTOPONÍMIA: ESTUDO DE CASOS

Maria Cândida Trindade Costa de Seabra

INTRODUÇÃO

Como desdobramento do Projeto Atlas Toponímico de Minas Gerais (ATEMIG), o subprojeto “Hidrotoponímia em Minas Gerais” constitui-se como uma pesquisa que abrange os estudos do homem e da sociedade por meio da linguagem e da investigação onomástica, ressaltando a interrelação língua, ambiente e cultura, ao dar destaque à água como elemento motivador no ato da denominação dos nomes de lugares.

O interesse na preservação dos nomes de lugares e o reconhecimento do papel relevante da natureza – em especial das águas em geral – no momento em que vivemos, levou-nos a pensar estratégias de valorização e revitalização dos “hidrotopônimos” que, de acordo com pesquisa já realizada, nos últimos anos no âmbito Projeto ATEMIG, constituem uma classe taxonômica bastante produtiva em Minas Gerais.

Caracterizando-se como um projeto regional, a proposta de estudar os hidrotopônimos – “topônimos resultantes de acidentes hidrográficos em geral”, como água, rio, córrego, ribeirão, braço, foz (Dick, 1990, p. 31) – presentes em território mineiro tem como objetivo mais amplo contribuir para o conhecimento e descrição da

realidade linguística do Estado de Minas Gerais. Essa descrição toponímica irá permitir ter uma visão ampla da nomeação nessa parte do território brasileiro, como: caracterização das línguas portuguesa, indígena, africana e outras; delimitação de áreas dialetais, identificação de mudanças em curso na língua, segundo variáveis diatópicas, além de favorecer a intercomparação de dados de natureza fonético-fonológica, lexical e morfossintática que costumam caracterizar o português.

Centra-se, portanto, em dois eixos, essa proposta de pesquisa:

a. dar continuidade às atividades inerentes ao Projeto ATEMIG sob nossa coordenação, na estrutura geral do Projeto (coleta de dados, armazenamento, transcrição dos dados e respectivas revisões, classificações etc.);

b. estudar de forma sistemática os hidrotopônimos, a partir dos dados documentados pelo Projeto ATEMIG, em especial o Banco de Dados escritos de base sincrônica (IBGE), em mapas históricos, cujos dados se encontram disponibilizados em <http://repositoriotoponomia.com.br/home>, em documentos escritos como registros de terras, cartas de sesmaria e, ainda, em entrevistas orais.

Este texto contempla o segundo item, ou seja, seleciona para estudo algumas denominações toponímicas presentes em nosso banco de dados, nomes de lugares que sintetizam a percepção do denominador em épocas pretéritas, quando “bons ares e boas águas” eram requisitos essenciais para a fixação de povoações no território brasileiro. Em um contexto singular, que evoca a exploração de um território, principalmente, a lida do homem com o gado e com a atividade minerária; a “água”, como, também, toda a rede semântica que a envolve torna-se, além de vital, utilitária.

1. A IMPORTÂNCIA DE ESTUDO DE CASOS EM TOPONÍMIA

As metodologias para análise de evidências empíricas na ciência linguística caracterizam-se, quase sempre, por serem qualitativas ou quantitativas, dependendo do objetivo da investigação e da escolha do pesquisador.

Ao contrário das pesquisas mais generalistas, os estudos de caso permitem a observação de detalhes da trajetória percorrida por um topônimo ao longo do tempo. Como características de “cases”, nunca a pesquisa foca em um grande grupo; nesta situação, em um conjunto toponímico, já que a proposta é sempre produzir conhecimento aprofundado sobre um nome específico – possui, portanto, caráter qualitativo, não quantitativo, conforme aponta Stake (1995, p.11), quando afirma que estudo de caso qualitativo é um “estudo da particularidade e complexidade de um único caso, chegando a compreender a sua atividade dentro de circunstâncias importantes”. Em síntese, o estudo de caso é um método de pesquisa que utiliza, geralmente, dados qualitativos, coletados a partir de eventos reais, com o objetivo de explicar, explorar ou descrever fenômenos atuais inseridos em seu próprio contexto. Caracteriza-se por ser um estudo detalhado e exaustivo de poucos, ou mesmo de um único objeto, fornecendo conhecimentos profundos (Eisenhardt, 1989; Yin, 2009).

Em se tratando do léxico toponímico, os estudos de caso são bastante salutares. Por se tratar de um referente, batizando um espaço geográfico, a palavra, em sua modalidade toponímica, reveste-se de caráter denominativo, referindo-se a um conjunto de descrições que chamaram a atenção do nomeador no passado, mas, muitas vezes, por não fazerem mais parte do vocabulário comum de membros de uma

comunidade linguística, essa palavra, com o passar do tempo, adquire uma opacidade semântica, sendo ignorado, assim, o sentido que motivara sua nomeação. No entanto, mesmo tendo o significado ignorado, o topônimo continua a exercer a função referencial.

O topônimo é, portanto, uma entidade que vai além da expressão linguística e envolve, obrigatoriamente, o referente que destaca. É um designador rígido. Mais que isso, os nomes de lugares, assim como os nomes de pessoas, representam ou são os próprios referentes. Logo, lhes é atribuído um estatuto diferente. Dentro dessa “teoria causal da referência”, Oliveira (1996, p. 351) diz que o nome próprio é um “designador rígido”, pois designa um indivíduo ou um lugar de uma maneira única e direta.

Sabendo que cada caso é único, essa modalidade de estudo de caso interessa aos estudos toponímicos, uma vez que, ao aprofundar a investigação sobre um nome de lugar, identifica-se sua trajetória e, muitas vezes, comportamentos relevantes do passado. Como primeiro passo, deve-se realizar uma revisão da literatura disponível, procurar coletar dados, familiarizar-se com o universo a ser investigado, consultar documentos e registros em arquivos, além de realizar entrevistas. Essas ferramentas, importantes para um estudo de caso qualitativo, possibilitam que o pesquisador amplie sua visão sobre o assunto ou o tema apresentado. Ao explorar esses estudos, criam-se fundamentos para a realização de pesquisas mais aprofundadas, que só poderão ser desenvolvidas quando o pesquisador tiver maior compreensão sobre o assunto estudado.

2. TOPÔNIMOS EM DESTAQUE

Damos início, neste texto, ao estudo de sete topônimos que se revestem de natureza hidronímica, mostrando a importância de cursos d'água na vida do homem. São topônimos identificados com nomes descritivos que recuperam aspectos do ambiente físico em que se encontram. Destacam-se: Vau, Córrego, Sumidouro, Salto, Bicas, Caldeirões, Água e suas variantes.

2.1. Vau

De origem latina, “vau”, inicialmente, designava o lugar do rio onde a água é pouco funda. Ao indicar essa característica local, esse nome geográfico, genérico ou geral alertava o homem de outrora que poderia atravessar para outra margem, a cavalo ou a pé, sem maiores riscos. Bluteau (1720, p. 360) define “vau” como “a paragem por onde se possa passar hum rio, ou hua lagoa, sem barco, sem nadar, & sem se enlodar”. Para Silva (1789, p. 508), “va’o, s. m. no rio, he o lugar onde elle he mais baixo, e se pôde vadear; *passar a vau*, vadear”.

Sobre a origem latina, evidencia Cunha (2010, p. 670):

vau sm. ‘baixio, trecho raso de rio, lago etc, que se pode passar a pé’ | *uaoo* XIII | Do lat. *Vādum* // **vadeAR** vb. ‘passar ou atravessar a vau’ XVI. Do cast. *vadear* || *vade·ÁVEL* || 1844 || **vadoso** adj. ‘onde há vau’ 1813. Do lat. *vadosus*.

É um nome característico do Brasil rural, sendo classificado como regional, conforme atesta Souza (2004, p. 335):

Aparecem, porém, no sertão do Brasil, certas expressões regionais como sejam, em Goiás, as seguintes que nos foram explicadas pelo prof.

Jubé: *vau de orelha, vau de cauda, vau a pés enxutos*. “O *vau de orelha* é a passagem que, através de um rio, só é possível com o animal a nado, dizendo-se de orelha por serem as orelhas as partes mais visíveis do animal quando a nadar; o *vau de cauda* é menos profundo e se diz assim quando o animal atravessando o rio as águas só atingem a cauda ou até a barriga; o *vau a pés enxutos* é uma variedade dos *vau*s para indicar aquele que, facilmente, pode ser atravessado, sem pisar-se na água, isto é, pulando-se de pedra em pedra, quando elas existem fora do nível das águas, o que é muito comum nos arroios e córregos de pequena largura” (Carta de 5-1-930).

Algumas dessas expressões já integravam o verbete “vao” em Bluteau (1720, p. 360):

Sondar o vao. *Tentare vadum. Ovid.* Também é usado no sentido figurado.

Passar o vao, *vid.* Vadear. (A qual ribeyra passei a *Vao*. Corographia de Barreiros, 179, ver.)

Adagios Portuguezes do Vao.

Por velho que seja o barco, sempre passa o vao,

Vao dorelha he perigoso.

Nem rio sem vao, nem geração sem mao.

Alto para vao, baixo para barco.

O dicionário *Aulete digital* aponta outras expressões, além das já citadas por Souza (2004):

A vau: A pé, dentro de rio (diz-se ger. de travessia de rio);

Dar vau: Permitir a passagem a vau (rio ou trecho de rio);

Fazer vau: Atravessar a vau na frente, como guia da travessia.

Além de acidente hidrográfico, consta “vau”, na língua portuguesa do Brasil e de Portugal, como nome próprio de lugar, embora, como já assinalado, figure, ainda, como nome genérico em algumas áreas rurais no sertão do Brasil. De acordo com Guerra (1987, p. 433), “vau é palavra portuguesa usada para os trechos de rio onde o nível das águas permite travessia a pé, ou a cavalo. Trata-se, por conseguinte, de um termo descritivo, usado principalmente pela geografia”.

Machado (1984, p. 1461) define *Vau* como topônimo frequente e mostra sua ancianidade: “Em 1258, ‘Valo de Veeiro (Inquisitiones, p. 319), Vao da Devesa (id., p. 293)”. Também, como nome próprio ou topônimo, o Banco de Dados do Projeto ATEMIG (FALE/UFMG) registra “vau” e suas variantes compostas e derivada, nomeando fazendas, córregos, ilhas, ribeirões, localidades, conforme apresenta-se a seguir:

Vau, Vau-Açu, Vau da Boiada, Vau da Ema, Vau da Lagoa, Vau da Saia, Vau das Éguas, Vau das Pedras, Vau do Cerradão, Vau do Palmital, Vau do Paraíso, Vau dos Patos, Vau Flor, Vau Grande, Vau Seco, Vau Verde, Santo Antônio do Vau Verde, Vauzinho.

Na forma simples, destacam-se os topônimos *Vau* e a forma derivada *Vauzinho*. As formas compostas apresentam-se com preposição e sem preposição:

a. Sem preposição:

Vau-Açu, Vau Grande, Vau Seco, Vau Verde, Vau Flor, (Santo Antônio do) Vau Verde.

b. Com preposição:

Vau da Ema, Vau das Êguas, Vau dos Patos, Vau da Boiada, Vau da Lagoa, Vau da Saia, Vau das Pedras, Vau do Cerradão, Vau do Palmital, Vau do Paraíso.

Referindo-se à dimensão do “vau”, compõe o nome os adjetivos *grande* e *açu*, em língua portuguesa e tupi, respectivamente, formando *Vau Grande* e *Vau Açú*. Ressalta-se, ainda, a predominância de compostos referentes à natureza física, em destaque: *Vau das Pedras, Vau do Cerradão, Vau do Palmital, Vau Flor*.

2.2. Córrego

O “Glossário de Termos Genéricos dos Nomes Geográficos utilizados no mapeamento sistemático do Brasil” (IBGE, 2010, p. 16) define “córrego” como “curso de água corrente de pequeno porte, que se encontra presente em todas as regiões fisiográficas brasileiras, na maioria das Unidades da Federação”.

Já as definições apontadas por dicionários *online* da língua portuguesa trazem acepções bem mais amplas para a unidade léxica *córrego* como as apresentadas pelo Michaelis Dicionário Brasileiro da Língua Portuguesa (2015). Nessa obra, *córrego* é assim definido:

- 1 Sulco aberto pelas águas correntes; corga, regueiro.
- 2 Pequeno rio, pouco profundo e de pequeno caudal; corgo.
- 3 Caminho estreito entre montes ou muros.
- 4 REG (BA) Denominação aos afluentes do rio São Francisco.

Como se pode destacar, *córrego*, além de designar conduto de água, indica “caminho estreito entre montes ou muros”, significado pouco conhecido entre os brasileiros, mesmo para aqueles que moram no campo. Contemplando essa terceira acepção dada por vários dicionários, dentre eles o dicionário Michaelis (2015), Lima (2021), em sua pesquisa de mestrado intitulada “A toponímia rural do contexto cafeeicultor da serra do Caparaó”, aponta vários nomes próprios de lugares que são chamados de *córregos*, sem, no entanto, terem como motivação a água e, sim, a geografia da região, referindo-se à sua forma topográfica, um “caminho estreito entre montes ou montanhas”; em evidência os topônimos “Córrego dos Lima”, “Córrego dos Tavares”, “Córrego dos Alves”, “Córrego Apolinária”, “Córrego Bragança”, “Córrego Vargem Alegre” etc, todos se referindo a bairros rurais.

Embora pouco conhecida, essa acepção é frequentemente destacada em dicionários contemporâneos de língua portuguesa, como já mostrado no dicionário Michaelis (2015) e destacado, na segunda acepção, em Houaiss (2001):

Córrego 1 fenda ou sulco aberto na terra pelas águas correntes; corga 2 via estreita e funda entre montes, cadeia de montanhas etc.; corgo, desfiladeiro 3. *B* pequeno rio com fluxo de água bastante tênue; corgo, riacho 4 *B N.E.* cada um dos afluentes do rio São Francisco. ETIM lat.hsp. *corrūgus*, ‘írego ou vala de água onde se lavavam metais’; f. hist.. 1087 *còrrago*, sXV *córrego* SIN/VAR arroio, corja, corjão, corjo, levada, regato, riacho, ribeira, ribeirada, ribeirão, ribeiro, sanga, veia, veio; ver tb. Sinonímia de *desfiladeiro* e *rio*.

Cunha (2010, p. 183) mostra a sua origem no latim **corrugus*, enquanto Machado (1984, p. 454) diz ser o vocábulo *córrego* “topônimo frequente no Brasil, na Galiza e em Portugal”. Acusam-se, no

Banco de Dados do Projeto ATEMIG, as variantes, *corgo*, *corgão* e *corguinho*, tanto como topônimo quanto como termo genérico. No estudo de Lima (2021) a forma “*corgo*” também se destaca como nome próprio de lugar.

Presente na toponímia mineira e no nordeste, conforme atesta Souza (2004), essa variante de *córrego* é bastante frequente, principalmente quando se trata do português em sua modalidade oral.

Corgo: palavra usada na linguagem popular portuguesa no sentido de riacho, *córrego*, *regueiro*; *regueira*. Com este significado é usada no Brasil, principalmente no Nordeste e em Minas Gerais segundo informa Aug. Saint-Hilaire. Designa também represa das águas do rio cheio em estreitos vales marginais: neste sentido ouvimo-lo no nordeste da Bahia (Souza, 2004, p. 117).

Souza (2004, p. 117) registra, ainda, *corguinho*, como diminutivo de *corgo*. Essa forma diminutiva é encontrada como topônimo no mapa de Eschwege (1821) nomeando freguesia, capela e arraial, na Comarca do Serro Frio em Minas Gerais (Santos; Seabra; Costa, 2017).

Além de *corgo*, *córrego*, *corguinho*, Machado (1984, p. 450) aponta a forma *corga*, como topônimo frequente em Portugal e na Galiza, “o mesmo que *corgo* (*córrego*); ‘em Mértola *corga* significa – encosta suave, menos íngreme que a ladeira’ –, segundo Leite de Vasconcelos [...]. *Corga* frequentes: na Galiza: Pontevedra”.

2.3. Sumidouro

“*Somjdoyro*”, século XV, derivado do verbo “sumir”, é topônimo presente no Brasil e em Portugal conforme atesta Machado (1984, p. 1372) que assim define o termo: “local por onde desaparece um curso de água que passa a ter leito subterrâneo”.

De vocabulário geral da língua, migrou para nome próprio de lugar, nomeando em território mineiro córregos, fazendas, localidades, cachoeira, povoado, vereda, morro, chapada, rio, povoado, lagoa. O Banco de Dados do Projeto ATEMIG atesta a forma “Sumidouro” e “Quinta do Sumidouro” como topônimos presentes em nossa contemporaneidade. Já o Banco de Dados históricos, disponível no “Repositório de Toponímia” apresenta o topônimo “Sumidouro” no mapa de Eschwege (1821), nomeando freguesias, capelas e arraiais e no mapa de Miranda (1804), nomeando arraial e freguesia, todos localizados hoje no município de Padre Viegas.

Nos primórdios de Minas Gerais, em 1711, “Sumidoiro” já era topônimo na região recém-descoberta de ouro, como atesta a Carta de Sesmaria “dada neste Arrayal do Ribeirão do Carmo aos 27 dias do mez de Fevereiro de 1711” hoje cidade de Mariana.

[...] Faço saber aos que esta minha Carta de Sesmaria virem que [...] Manuel Antunes de Lemos que elle Suppte. pesuhia hú çitio no destrito deste Ribeirão de N. Sr^a do Carmo junto ao morro de Domingos Velho Cabral o qual comprou a André Lopes de Azevedo; e porque se achava com bastantes escravos p^a poder cultivar o d^o çitio, e neçcitava de mais terras para poder cultivar, e pagar dízimos a Deus; portanto me pedia fose servido conceder-lhe por Carta de Sismaria o d^o çitio com meya legoa de terras de testada correndo p^a a banda do Sumidoiro [...] Dada neste Arrayal do Ribeirão do Carmo aos 27 dias do mez de Fevereiro de 1711.¹

Com o fim da atividade minerária nessa região de Minas, houve apagamento do significado do nome, que era atrelado à lida com o ouro. De Sumidouro ~ Sumidoiro, o povoado, quando se tornou distrito, recebeu o nome de Padre Viegas.

¹ Revista do Arquivo Público Mineiro, ano 11, fascículo 2, 1897, p. 264-265.

Sumidouro (N. Sra. do Rosário) – Freguesia antiga e, outrora, das mais importantes da comarca de Mariana. Aqui funcionou por alguns anos o afamado colégio dos Padres Osórios, donde saíram homens notáveis, como o Padre Viegas, o Cônego Santa Apolônia, o Dr. Vicente Coelho de Seabra e muitos outros. Sumidouro é hoje freguesia anexa ao curato da Sé, donde dista nove quilômetros. A velha freguesia perdeu tudo, até o nome tradicional e glorioso. Hoje se chama Padre Viegas (Trindade, 1955, p. 303).

Na memória das pessoas mais velhas, como o relato do informante de 85 anos entrevistado por Seabra (2004, p. 292), transcrito a seguir, pode-se observar que o nome ainda é reconhecido como um termo geográfico, contudo, hoje, não se pode mais confirmar essa informação entre os mais jovens para quem o termo se tornou opaco.

Inf.: Sumidouro... acho que é... como é que a gente fala... quando o rio afunila e forma um poço que é perigoso então o povo chama de Sumidoro... Ali... naquela região... eles mexia com ouro... né? Mineração e acredito que deva ter algum lugar perigoso.

Pesq.: Mas as pessoas daquela região ainda falam Sumidouro... não falam Padre Viegas?

Inf.: Tem muito pouca gente... Os mais novos é que falam Padre Viegas... os antigos Sumidoro.

Pesq.: Por que isso heim?

Inf.: Ah... mantém o nome antigo né?

Sumidouro, como um termo que integra o vocabulário geral da língua portuguesa, teve seu significado ampliado, ligado, principalmente, a redes de esgoto nas grandes cidades brasileiras. Caracteriza-se como o ponto em que um curso de água superficial penetra no

solo, isto é, uma abertura natural que se comunica com uma rede de galerias e através da qual o curso de água adentra o subsolo.

2.4. Salto

Do latim “saltus –ūs”, *salto* é a denominação genérica dada a todos os tipos de desnivelamento ou degraus encontrados no perfil longitudinal de um rio; exemplo: cascata, catadupa, queda-d’água, cachoeira, corredeira, etc. [...] uma queda súbita das águas de um rio, como que havendo uma separação das partes: superior da inferior. As corredeiras e cascatas são formadas por fracos desnivelamentos em relação às cataratas, cachoeiras e quedas d’água (Guerra, 1987, p. 383).

Houaiss (2001) dedica duas entradas para *salto*. Na primeira entrada, na quarta acepção, define *salto* como “mesmo que queda d’água”; já a segunda entrada traz a marca de uso “obsoleto”. Sobre o étimo, destaca:

Salto [...] há controvérsia entre os etimologistas do latim sobre o vocábulo ‘saltus’. Ernout e Meillet consideram que, seja qual for seu significado, a palavra é uma só e se liga ao radical de saltum, supino de salire ‘saltar, pular’; para eles, da acepção de ‘salto, pulo’ passou-se às acepções de ‘passagem estreita, passo, desfiladeiro’, e, por estes lugares estarem sempre cobertos de vegetação, surgem as acepções ‘bosque, pasto’ e até mesmo a acepção de ‘medida agrária’; outros, entre os quais se destacam M. Lubke, Freund, Lewis e Short, Walde e Hofmann, e Glare, separam ‘saltus’ ‘pulo, salto’ de ‘saltus’ bosque, pastagem, desfiladeiro etc, ligando o primeiro ao verbo salire ‘saltar’, o segundo alguns como Lewis e também Glare consideram de origem duvidosa, não obstante associado às vezes a ‘saltus’ a forma ‘souto’, a datação é para a acepção ‘desfiladeiro’ (Houaiss, 2001, p. 2504).

Machado (1984, p. 1367) aponta *salto* como topônimo frequente em Portugal, na Galiza e no Brasil, no entanto, não relaciona esse nome a quedas d'água e, sim, a “lugar eminente; cerro, outeiro”. Esse autor não acredita que sua origem esteja diretamente ligada ao latim *saltu-*, que originaria “*Souto* – topônimo frequente em Portugal e na Galiza; aqui também ocorre (Lugo, Orense) com a grafia *Soto* (castelhanismo). Do s. m. *souto*. Formas antigas: *Sauto de fontano* em 906; *Sauto* (e variantes) em 950 [...]”.

Senna (1926, p. 282-283) mostra que em Minas as cachoeiras altas e volumosas são chamadas tombos, já um tombo forte e único constitui o Salto. Salto é o mesmo que “queda d'água”, podendo, ainda, ser sinônimo de mata, bosque, significado, segundo ele, um tanto arcaico, concordando com Houaiss (2001).

Além de nomear lugares, é um termo genérico muito comum no Brasil, sendo empregado, principalmente, em áreas rurais. Como topônimo, o Banco de Dados do Projeto ATEMIG atesta *salto* em formas simples e compostas, nomeando fazenda, localidade, municípios, rio, riacho, vereda, povoado, sítio, ribeirão, fazenda, córrego, a saber: *Salto, Buriti do Salto, Salto da Divisa, Santa Maria do Salto, Salto da Cabaça, Ribeirão do Salto, Salto do Ouro, Três Saltos, Salto de Cima, Salto do Meio, São Joaquim do Salto, Salto do Morais, Santa Cruz do Salto, Salto da Ponte*.

Como topônimo, encontramos *Salto* em documentos relativos a escrituras de terra, como a citada a seguir:

[...] pelos outorgantes vendedores me foi dito que são senhores e legítimos possuidores de parte do imóvel Salazar, deste município, com uma área de quatro hectares cinco ares e trinta e cinco centiares (4,05,35) de terras de cultura, com as seguintes confrontações: Partindo de um pau de cedro, no fundo,

aí pelo correjo acima até um toco de lei nova, deste em rumo direito a uma moita de taquarassú ao alto do **Salto**, pelo espigão abaixo com José Lana Junior até o rumo do pau de cedro, ponto de partida, confrontando com herdeiros de Ribeiro de tal, e com herdeiros de Severina de tal.[...] Barra Longa, 7 de março de 1955 (Seabra, 2004, p. 284).

2.5. Bicas

Do latim *beccus*, bica no sentido de torneira, “canudo, por onde sahe a agoa da fonte” (Bluteau, 1720), tem sua origem no século XIV. Segundo Nelson de Senna (1926, p. 252),

[...] em Minas, dá-se o nome de ‘bicas’ ou bicâmes aos conductos d’agua, feitos de madeira: bicas de cascas de certas palmeiras indigenas; bicas previamente lavradas (bicâmes de braúna, por exemplo). [...] Entre nós é corrente dizer: a bica d’água; o bicâme do engenho; as bicas de palmito (sempre no sentido de taes peças conductoras do precioso liquido, ou desses aquedutos rusticos feitos de madeira). A bica é sempre descoberta, o bicâme costuma ser tampado.

Machado (1984, p. 253) destaca *Bica* como antropônimo e como topônimo:

Bica¹, *apel.* [...] Do s.f. *bica*, “canal donde sai água”. Antiga alcunha, portanto, que se dava a quem possuísse ou morasse perto de *bica*. [...]

Bica², *top.* muito frequente, em nomes simples ou compostos (como *Chão de Bica*, *Mação*, *Pisão da Bica*, *Gouveia*, *Ponte da Bica*, *Paredes*, etc.) Do s.f. *bica*, “canal donde sai água”. [...] Também há *Bicas*, não tão frequente.

Como topônimo, presente em mapas históricos de Minas Gerais, é apontada a forma *Bicas*. Essa forma pluralizada consta nos mapas de Eschwege (1821) e de Miranda (1804), nomeando freguesias, capelas e arraiais no município de São Joaquim de Bicas.

O Projeto ATEMIG atesta a permanência desse nome na contemporaneidade, como topônimo presente em território mineiro, nomeando acidentes físicos e humanos, em destaque:

a. nomeando acidentes físicos: córrego da *Bica*, córrego das *Bicas*, córrego das *Três Bicas* (município de Oliveira), serras das *Bicas*, morro das *Bicas*, ribeirão das *Bicas*, rio de *Bicas*;

b. nomeando acidentes humanos: fazenda das *Bicas*, fazenda *Bicas de Cima*, povoado *Bicas*, município *São Joaquim de Bicas*, município *Bicas*, município *Biquinhas*.

Constata-se que há predominância da forma flexionada *Bicas*. A forma *Bica* só foi contabilizada uma vez. Em se tratando da forma derivada *Bicame* (*bica*+sufixo *-ame*), o Banco de Dados do Projeto ATEMIG registra o topônimo,

a. nomeando acidentes físicos: córrego *Bicâme* (município de Felício dos Santos, na região do Jequitinhonha), córrego *Bicame* (município de Santo Antônio do Amparo), cachoeira *Bicame* (Santana do Riacho), Cachoeira do *Bicame* (Lapinha da Serra),

b. nomeando acidentes humanos: fazenda *Bicâme* (município de São Roque de Minas).

2.6. Caldeirões

Como termo da natureza física, usado sempre na forma flexionada plural, *Caldeirões* costuma ser registrado em dicionários com várias acepções, de acordo com Souza (2004, p. 70):

Caldeirão: várias são as acepções deste termo. Na Amazônia significa remoinhos dos rios, isto é, lugares nos leitos dos rios onde se formam vórtices e turbilhões resultantes de correntes circulares, perigosas à navegação [...]. Na Bahia e região das Lavras Diamantinas assim se denominam buracos redondos cheios de cascalho, na piçarra das catas; ao caldeirão pequeno chamam *casco de burro*. No Rio Grande do Sul, chamam *caldeirão* a uma grande escavação no meio do campo, ou das estradas feitas pelas chuvas ou pisadas dos animais. Assim também em São Paulo [...]. Em outras zonas, assim se nomeiam tanques naturais nos lajedos, onde se armazenam águas pluviais. [...] Na *Geologia* de J. C. Braner, à pág. 28, primeira edição, encontramos *caldeirão* no sentido de cova lisa e arredondada, de várias dimensões, que se forma no leito dos rios, devida ao desgastamento produzido pelos remoinhos nas correntes, onde as pedras soltas são revolvidas em um círculo de pequeno diâmetro. [...] São mais frequentes junto a cascatas e quedas d'água e alguns os denominam *marmitas-de-gigante*, principalmente quando a profundidade é muito maior que a largura. Beaurepaire Rohan, registrando o termo *caldeirão*, dá-lhe o sentido de covas atoladiças, que se formam transversal e paralelamente, nas estradas frequentadas por tropas de animais no tempo das chuvas e acrescenta que, em Pernambuco e Alagoas, chamam a isso *camaleões*.

Em Minas Gerais, o topônimo *Caldeirões*, em regiões de lavras, está relacionado à atividade minerária, conforme relata Nelson de Senna (1924):

[...] no município de Ouro Preto, no districto de São Domingos, existe a antiga fazenda dos '*Caldeirões*'. E' um termo brasileiro mais empregado no plural e com peculiar significado. – Quasi sempre na superfície do arenite e logo abaixo da soleira de um salto, nas margens de rios encachoeirados, se observam muitas cavidades, cheias de seixos, que não são mais do que '*caldeirões*' ou pilões cavados na rocha pelas aguas, na época em que a soleira do salto se achava mais abaixo do lugar actual. O phenomeno é muito comum nos rios mineiros, desde o Rio Grande ao Jequitinhonha e mórmente na região da Serra do Espinhaço. O nome de '*caldeirão*' é, pois, um brasileirismo e já acceto pela sciencia, no sentido em que em Minas o empregamos (Seabra, 2004, p. 152).

Em um dos relatos do *Códice Costa Matoso*, registra-se:

Modo de extração dos cascalhos: Nos rios grandes, é necessário tirá-los da sua natural corrente, passando-os, ou por bicas ou por terra, rompendo-se em alguma parte, conforme se precisa e é mais conveniente. E mudando assim o rio, se lhe assenta uma roda de esgotar água por rosário, e esgotada se vai tirando o desmonte ou entulho da terra, por baixo do qual está o cascalho que se tira à bateia para ser lavado na forma acima. E em partes do rio Jequitinhonha se lhe acham poços, a que chamam *caldeirões*, de 120 e 140 palmos de fundo, tendo dentro em si várias lapas e concavidades que metem para as serras que estão nos lados do rio imediatas a ele. (Modo de extração dos cascalhos. Anônimo, Vila do Príncipe; ca.1750, *apud* Lima, 1999, p. 850).

Sobre a importância dos caldeirões, relacionada à atividade da mineração, salienta Paul Ferrand (1998, p. 95-96):

Os depósitos de aluviões se formaram nos vales ou nos leitos dos rios; as pedras e o ouro arrancados dos flancos das montanhas pelas águas se acumularam nos fundos dos vales, formando um leito de seixos rolados, de areia e de ouro, designado na região como ‘cascalho’, cujo enriquecimento se produziu por uma espécie de concentração mecânica natural. Os cursos de água, ao penetrar nas fraturas, escavam, pouco a pouco, cavidades que apresentavam o aspecto de enormes caldeiras, com forma cilíndrica. O fundo, em meia esfera e as paredes, ficam lisos e polidos pela rotação da água que arrastavam consigo seixos e fragmentos de rochas. Essas cavidades, as quais se dá o nome de ‘caldeirões’, encheram-se, com o tempo, de matéria mais pesada formando depósitos de cascalho aurífero de riqueza frequentemente saliente.

Em Portugal, José Pedro Machado (1984, p. 315) apresenta *Caldeirões* como topônimo em várias localidades, dentre elas, “Alandroal (Monte dos Caldeirões), Covilhã, Santiago do Cacém, na Ilha Terceira”.

2.7. Água

Em se tratando de topônimos, é consensual afirmar que o trio – motivação, convenção, identificação – são três constantes fundamentais no ato de nomear, apoiadas na estruturação léxica. Devido à importância dos cursos d’água para a vida das pessoas – principalmente, para sua subsistência, locomoção e trabalho – pode-se comprovar sua importância, também, quando surgem os nomes de lugares revestidos de natureza hídrica. Seja indicando caminhos, seja apontando

os locais onde se encontra o precioso mineral, seja identificando a qualidade da água ou, ainda, caracterizando-a, esse elemento da natureza deixa rastros, que são encontrados nos nomes de lugares, em acidentes físicos, como rios, córregos e em acidentes humanos, como fazendas e cidades.

De natureza hídrica, impressiona o número de topônimos compostos a partir do elemento água. Formados por meio do processo conhecido por composição, que se caracteriza por ser o processo em que melhor se visualiza o espírito criador da língua, revelando formações que denotam um sentido agudo de observação e de expressividade, com características bastante populares; destacam-se inúmeros topônimos que se repetem ao longo de todo o território mineiro. Na maioria das vezes, compõe vocábulos transparentes, dispensando comentários. Tendo como destaque a água, o Banco de Dados do Projeto ATEMIG guarda um número muito grande desses compostos, nomeando córregos, rios, riachos, fazendas, sítios, lagoas, povoados, municípios no estado de Minas Gerais. Em relação ao aspecto formal desses topônimos, convém distinguir:

a. substantivo + adjetivo: a composição faz-se por coordenação de um adjetivo a um substantivo. *Água Amarela, Água Azul, Água Barrenta, Água Bela, Água Boa, Água Bonita, Água Branca, Água Catinguenta, Água Choca, Água Clara, Água Doce, Água Emendada, Água Escura, Água Espalhada, Água Esparramada, Água Espriada, Água Feia, Água Férrea, Água Formosa, Água Fria, Água Funda, Água Furtada, Água Limpa, Água Limpinha, Água Luminosa, Água Mansa, Água Mineral, Água Morena, Água Nova, Água Parada, Água Podre, Água Presa, Água Preta, Água Quente, Água Rasa, Água Rica, Água Ruim, Água Salgada, Água Salobra, Água Santa, Água Sempre-Viva, Água Sumida, Água*

Suja, Água Tupiara, Água Verde, Água Vermelha, Água Virada, Água Viva.

Desses topônimos, merecem comentários por não terem significados tão conhecidos:

- *Água Emendada* – nascente ou desaguadouro comum de dois ou mais rios pertencentes a bacias diferentes (Souza, 2004, p. 05).

- *Água Espraiada* – derivada de praia, *espraiar* integra o léxico da língua portuguesa desde o século XV. “Espraiado”, segundo Souza (2014, p. 141) pode se referir à “expansão de um rio, alargamento de seu leito, quase sempre de pouca profundidade e margens arenosas” ou “ribeirão que corre em leito raso, geralmente de areia”, definição tomada por ele de Amadeu Amaral.

b. substantivo + adjetivo (flexionados no plural): o número de ocorrências é bem menor. *Águas Altas, Águas Belas, Águas Claras, Águas Pretas, Águas Tiradas, Olhos D’água;*

c. adjetivo + substantivo: nesta ordem o número de ocorrências é menor do que no item “a”. *Muita Água, Meia Água.*

d. verbo + substantivo: para formar nomes compostos como esse, utiliza-se, comumente, a terceira pessoa do singular, do modo indicativo ou do imperativo. Coletamos o topônimo composto *Bebe-Água.*

e. substantivo + de + substantivo: também este tipo de composição, em que um substantivo está subordinado a outro, está

bem representado nos dados analisados. *Água da Serra, Água de Arroz, Água da Estaca, Água de Contendas, Água do Campo, Água do Pastoril, Água do Prata, Água dos Veados, Capão D'água, Capão da Água, Córrego D'água, Lapa D'água, Mãe D'água, Olho D'água, Pedra D'água, Pingo d'água, Poço D'água, Ponta D'água, Rego D'água, Riacho D'água, Tirada D'água, Veio d'água, Vereda D'água, Vereda das Águas.*

f. substantivo + substantivo + de + substantivo: *Córrego Olho D'Água.*

g. substantivo + de + substantivo + adjetivo: *Olho D'água Comprida, Barra da Água Suja, Cabeceira da Água Santa, Água do Morro D'água.*

h. substantivo + de + substantivo + de + substantivo: *Saco dos Olhos d'Água, Olho D'água das Minhocas, Buriti do Olho d'água.*

i. substantivo + de + substantivo + de + substantivo + adjetivo: *Olhos D'água do Barro Vermelho.*

j. substantivo + adjetivo + de + substantivo + adjetivo: *Água Limpa do Mato Virgem.*

k. substantivo + adjetivo + de + substantivo: *Água Emendada da Serra, Água Limpa da Serra.*

Da mesma natureza, dentre as formações simples e derivadas que contemplam o elemento água, constam:

a. *Água*: poucas ocorrências.

b. *Águas*: ocorre um pouco mais, se comparada a forma singular.

c. *Aguada*: forma derivada, formada pelo processo de derivação lexical.

Combinando o princípio de derivação lexical (água + sufixo -ada), seguido de composição, tem-se os topônimos *Capão da Aguada*, *Aguada Nova*.

De acordo com Weinreich (1953, p. 54), “a necessidade de designar novas realidades, fenômenos, lugares e conceitos é, obviamente, a causa universal da inovação lexical”. No caso de topônimos, a composição e a derivação, ou os dois processos juntos, constituem uma fonte quase inesgotável de inovação lexical, ampliando e, ao mesmo tempo, particularizando o acidente nomeado.

Atualmente, nomeando municípios, o estado de Minas Gerais conta com cinco topônimos, que têm como motivação o elemento água, a saber: *Água Boa*, *Água Comprida*, *Aguanil*, *Águas Formosas*, *Águas Vermelhas* – todos formados pelo processo de composição – tendo em *Aguanil*, a composição por aglutinação (água + anil, referente à cor azul). No passado, esse número era bem maior, ocorrendo as seguintes substituições:

a. *Água Limpa* > *Aguanil*;

b. *Água Suja* > *Água Limpa* > *Berilo*;

c. *Água Limpa* > *Coronel Pacheco*;

d. *Água Quente* > *Águas Quentes* > *Santana da Água Quente* > *Montezuma*;

e. *Água Santa* > *Águas Santas* > *Águas Virtuosas* > *Águas Virtuosas da Campanha* > *Lambari*;

f. *Águas Virtuosas de Caldas > Nossa Senhora da Saúde das Águas de Caldas > Nossa Senhora da Saúde de Caldas > Nossa Senhora da Saúde de Poços de Caldas > Poços de Caldas;*

g. *Águas Virtuosas de Cambuquira > São Sebastião de Cambuquira > Cambuquira.*

CONSIDERAÇÕES FINAIS

Sabemos que os estudos toponímicos, motivados pelo elemento hídrico em Minas Gerais, não se encerram com esses sete casos apresentados – há um número de dados muito maior que contempla o tema relacionado à água. Se, por um lado, o Projeto ATEMIG conta com dados já sistematizados, e, ainda, com estudos pontuais sobre a hidrotoponímia e a hidronímia, por outro, há que se realizar uma “colheita” muito mais rica de vestígios do passado que são os nomes que poderemos encontrar na microtoponímia. Além disso, os mapas históricos continuam sendo objeto de estudos e deles devem advir novas formas toponímicas, contribuindo desse modo com as pesquisas relacionadas à variação e à mudança.

Há muito que se fazer. O homem ao nomear um lugar amplia e ao mesmo tempo singulariza aquele espaço, por isso, por serem eminentemente singulares, os topônimos são muitas vezes complexos, escapando a toda investigação. Percebemos variações regionais dentro do próprio estado; notamos, também, que, muitas vezes, em um ambiente, uma linha tênue separa as taxonomias, o que requer que aprofundemos, cada vez mais, estudando casos, cuidando para que não se perca, na análise a carga cultural do nome.

REFERÊNCIAS

AULETE, Caldas. **Aulete Digital** – Dicionário contemporâneo da língua portuguesa. São Paulo; Melhoramentos. Disponível em: <https://www.aulete.com.br/>. Acesso em: 14 out. 2020.

BLUTEAU, Rafael. **Vocabulário Português e latino** (v. 7: Letras Q-S). Lisboa: Oficina de Pascoal da Sylva, Impressor de Sua Magestade, 1720. Disponível em: <https://digital.bbm.usp.br/handle/bbm/5443>. Acesso em: 15 nov. 2020.

CUNHA, Antônio Geraldo da. **Dicionário Etimológico da Língua Portuguesa**. Rio de Janeiro: Lexicon, 2010.

DICK, Maria Vicentina de Paula do Amaral. **Toponímia e Antroponímia no Brasil**. Coletânea de Estudos. 2. ed. São Paulo: FFLCH/USP, 1990.

EISENHARDT, Kathleen Marie. Building Theories from Case Study Research. **Academy of Management Review**, v. 14, n. 4, p. 532-550, 1989. Disponível em: https://josephmahoney.web.illinois.edu/BADM504_Fall%202019/Eisenhardt1989.pdf. Acesso em: 22 abr. 2021.

FERRAND, Paul. **O ouro em Minas Gerais**. Belo Horizonte: Fundação João Pinheiro, 1998.

GUERRA, Antônio Teixeira. **Dicionário geológico-geomorfológico**. 7. ed. Rio de Janeiro: IBGE, 1987.

HOUAISS, Antônio. **Dicionário Houaiss da Língua Portuguesa**. Rio de Janeiro: Objetiva, 2001.

IBGE - Instituto Brasileiro de Geografia e Estatística. **Glossário dos termos genéricos dos nomes geográficos utilizados no mapeamento sistemático do Brasil**. Rio de Janeiro: IBGE, 2010. Disponível em: <https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?id=288835&view=detalhes>. Acesso em: 13 nov. 2020.

LIMA, Jacqueline Helen. **A toponímia rural no contexto cafeeicultor da Serra do Caparaó**. 2021. 169 f. Dissertação (Mestrado em Linguística Teórica e Descritiva) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2021.

LIMA, Yêdda Dias. **Código Costa Matoso**: Coleção das notícias dos primeiros descobrimentos das minas na América que fez o doutor Caetano da Costa

Matoso sendo ouvidor-geral das do Ouro Preto, de que tomou posse em fevereiro de 1749, & vários papéis. Belo Horizonte: Fundação João Pinheiro, Centro de Estudos Históricos e Culturais, 1999. 2v.: il. mapas (Coleção Mineiriana, Série Obras de Referência).

MACHADO, José Pedro. **Dicionário Onomástico e Etimológico da Língua Portuguesa**. Lisboa: Editorial Confluência, 1984. v. 3.

MICHAELIS. **Moderno dicionário da língua portuguesa**. São Paulo: Melhoramentos, 2015. Disponível em: <https://michaelis.uol.com.br/moderno-portugues/>. Acesso em: 12 nov. 2020.

OLIVEIRA, Fátima. Semântica. *In*: FARIA, Isabel Hub; PEDRO, Emília Ribeiro; DUARTE, Inês; GOUVEIA, Carlos A. M. (orgs.). **Introdução à Linguística Geral e Portuguesa**. Lisboa: Editorial Caminho, 1996. p. 332-382.

REVISTA DO ARQUIVO PÚBLICO MINEIRO, v. II, fac. 2º, 1897.

SANTOS, Márcia Maria Duarte dos; SEABRA, Maria Cândida Trindade de; COSTA; Antônio Gilberto (Orgs.). **Toponímia Histórica de Minas Gerais**: Registros em Mapas do Setecentos ao Oitocentos Joanino. 1. ed. Belo Horizonte, MG: Museu de História Natural e Jardim Botânico, Universidade Federal de Minas Gerais (MHNJB-UFMG); Instituto Histórico e Geográfico de Minas Gerais (IHG-MG), 2017. Disponível em: <http://www.repositoriotoponomia.com.br.html>. Acesso em: 10 nov. 2020.

SEABRA, Maria Cândida Trindade Costa de. **A formação e a fixação da língua portuguesa em Minas Gerais**: a Toponímia da Região do Carmo. 2004, 368f. Tese (Doutorado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, 2004.

SEABRA, Maria Cândida Trindade Costa de. Pesquisa Toponímica em Minas Gerais: contribuições do Projeto ATEMIG. *In*: SOUZA, Olímpia Maluf; SILVA, Valdir; ALMEIDA, Eliana de; SALOMÃO, Leila; BISINOTO, Jacob (Orgs.). **Discurso, Sujeito e Memória**. Campinas: Pontes Editores, 2012. p. 305-315.

SEABRA, Maria Cândida Trindade Costa de. **Base de Dados do Projeto Atlas Toponímico do Estado de Minas Gerais**. Belo Horizonte: FALE/UFMG, 2018.

SENNA, Nelson de. Toponymia Geographica de origem brasílico-indígena em Minas Geraes. **Revista do Arquivo Público Mineiro**, Belo Horizonte, p. 191-337, 1924.

SENNA, Nelson de. Nótulas sobre a toponímia geográfica brasílico-indígena em Minas Gerais. **Revista do Arquivo Público Mineiro**, v. 20, p. 194-337, 1926.

SILVA, Antônio de Moraes, 1755-1824; BLUTEAU, Rafael, 1638-1734. **Dicionário da lingua portugueza composto pelo padre D. Rafael Bluteau, reformado, e accrescentado por Antonio de Moraes Silva natural do Rio de Janeiro** (v. 2: L - Z). Lisboa: Na Officina de Simão Thaddeo Ferreira, 1789. Disponível em: <https://digital.bbm.usp.br/handle/bbm/5413>. Acesso em: 12 jan. 2021.

SOUZA, Bernardino José. **Dicionário da terra e da gente do Brasil**. Belo Horizonte: Editora Itatiaia, 2004.

STAKE, Robert E. **The art of case study research**. Thousand Oaks: Sage Publications, 1995.

TRINDADE, Cônego Raimundo Octávio da. **Velhos Troncos Mineiros**. São Paulo: Revista dos Tribunais, 1955. v. 3.

WEINREICH, Uriel. **Language in contact**. New York: Linguistic Circle; The Hague: Mouton, 1953.

YIN, Robert. K. **Case study research: Design and methods**. 4. ed. Thousand Oaks, CA: Sage, 2009.

EXPERIÊNCIA DO ENSINO REMOTO DE LEXICOLOGIA E LEXICOGRAFIA

Rosemary Irene Castañeda Zanette

INTRODUÇÃO

Neste século, 2020 foi um dos anos mais marcantes. O mundo todo foi impactado pelo início da pandemia do novo Coronavírus (Sars-Cov-2), que ainda em 2021 está presente no cotidiano. Todos os setores da sociedade foram afetados por um problema que, a princípio, seria circunscrito ao âmbito da saúde. Medidas emergenciais foram tomadas com base em quando o vírus chegou a cada país e no agravamento da doença, popularmente conhecida como COVID-19. No Brasil, já se passou mais de um ano desde a confirmação do primeiro caso. No início de 2021, diferentes vacinas começaram a ser exportadas e outras produzidas em território nacional. E mesmo com o início de sua subministração, em várias faixas da população, de primeira e de segundas doses (e da possibilidade de terceiras), ainda se está distante de atingir um patamar suficiente de imunização que permita aos brasileiros retomarem as atividades sem tantas restrições, tendo em conta as diversas variantes da doença que surgiram.

No que tange à Educação, os reflexos desse momento crítico também persistem em 2021. Foi necessária uma reorganização das

atividades, em todo o país, desde a da Educação Básica ao Ensino Superior, dos cursos livres, enfim, de todo o setor. Nessa perspectiva, algumas diretrizes foram estabelecidas em âmbitos federal, estadual e municipal, bem como nas próprias instituições, conforme as particularidades de cada região ou localidade. Inicialmente, ainda em 2020, todas as atividades foram suspensas, no intuito de conter a transmissão do vírus. Sem a perspectiva de um breve retorno, foi necessário rever as medidas tomadas a fim de dar prosseguimento às atividades e minimizar os prejuízos aos principais envolvidos, alunos e professores e demais profissionais do setor. Surgiram duas opções: a do ensino à distância e a do ensino remoto emergencial, aplicados a depender das condições de cada nível de ensino e de cada região ou localidade.

Apesar de, em 2021, a pandemia estar em outro momento, com medidas sanitárias que permitem o retorno das atividades com algumas restrições, bem como haver um Programa Nacional de Imunização (PNI) em andamento, neste trabalho, concentramo-nos em atividades realizadas ainda em 2020.

No Estado do Paraná, com relação ao Ensino Superior, gradativamente as instituições foram se posicionando sobre a situação. As medidas adotadas se apresentam iguais ou distintas para a Graduação e para a Pós-Graduação. Com relação a esta última, a Universidade Estadual do Oeste do Paraná (Unioeste) tomou uma decisão mais tardia em relação às demais universidades estaduais. Em maio, com base na Resolução Nº 52/2020-CEPE, foi autorizada a substituição das aulas presenciais pelas aulas remotas síncronas. Em seguida, cada Programa de Pós-Graduação teve de aprovar seu novo calendário acadêmico. Houve também uma nova chamada para as matrículas em disciplinas, e os discentes, ao se matricularem, deveriam assinar um Termo de Ciência indicando que eram dotados de recursos tecnológicos e condições para participar das aulas nesse novo formato.

No caso do Programa de Pós-Graduação em Letras, sediado na Unioeste, *campus* de Cascavel (PR), ofertamos, em 2020, a disciplina de *Lexicologia e Lexicografia*, na modalidade remota, com atividades síncronas e assíncronas, no período de 07 de julho a 01 de setembro de 2020. Realizamos encontros síncronos, por meio da plataforma *Google Meet*, com duração de 1h30min, e as demais atividades assíncronas, antes e após cada aula, realizadas no *Google Sala de Aula*. Inscreveram-se 17 alunos regulares e 7 alunos especiais, mas houve desistências de alunos especiais e de um regular; o grupo final foi de 21 participantes.

Assim, neste capítulo, nosso escopo é refletir sobre o andamento e a realização dessa disciplina nesse novo formato. Para tanto, descrevemos e analisamos os três momentos da atividade: (i) a preparação; (ii) o desenvolvimento; (iii) a avaliação da disciplina. Este texto contempla, inicialmente, as reflexões teóricas que sustentaram a disciplina e, posteriormente, a descrição e a reflexão das ações desenvolvidas.

1. NORMATIVA E EMBASAMENTO TEÓRICO

No Brasil, a tomada de decisão quanto à realização das atividades de forma não presencial se deu após o entendimento de que não havia condições sanitárias seguras para que as atividades retomassem nesse formato. Pode-se dizer que esse foi um momento de incerteza, em que não se sabia muito como restringir o alcance do vírus, devido à ausência de tratamento preventivo e/ou imunização e de um tratamento eficaz das pessoas acometidas pelo novo Coronavírus. Nesse contexto, as diversas instituições relacionadas à Educação tiveram de se posicionar, determinando como a situação deveria ser gerenciada no setor. Em seguida, destacamos algumas normativas oficiais relati-

vas aos procedimentos que deveriam ser adotados no Ensino Superior durante o período da pandemia.

O primeiro documento importante elaborado pelo Ministério da Educação (MEC) foi a Portaria nº 343, de 17 de março de 2020 (Brasil, 2020a), emitida logo após a identificação dos primeiros casos de contaminação pelo vírus no país. O primeiro artigo dessa portaria apresentou o seguinte posicionamento:

Art. 1º Autorizar, em caráter excepcional, a substituição das disciplinas presenciais, em andamento, por aulas que utilizem meios e tecnologias de informação e comunicação, nos limites estabelecidos pela legislação em vigor, por instituição de educação superior integrante do sistema federal de ensino, de que trata o art. 2º do Decreto nº 9.235, de 15 de dezembro de 2017 (Brasil, 2020a).

A princípio, previu-se uma prorrogação inicial de 30 dias ou do período necessário, conforme orientações do Ministério da Saúde (MS). A portaria se refere às instituições federais, no entanto, o modelo foi seguido por outras instituições estaduais e/ou particulares.

Outro documento significativo foi o Decreto Legislativo nº 6 de 2020, de 20 de março de 2020 (Brasil, 2020b), que reconheceu o estado de calamidade pública instalado no país, cujo prazo foi determinado até o dia 31 de dezembro de 2020. Embora a denominação seja impactante, ela se traduz na possibilidade de alteração de uso do orçamento, previamente aprovado, para destiná-lo a outros fins, como o de assistência social.

Posteriormente, foi emitida também pelo Governo Federal a Portaria nº 544, de 16 de junho de 2020 (Brasil, 2020c), referente ao período em que durar a pandemia:

Art. 1º Autorizar, em caráter excepcional, a substituição das disciplinas presenciais, em cursos regularmente autorizados, por atividades letivas que utilizem recursos educacionais digitais, tecnologias de informação e comunicação ou outros meios convencionais, por instituição de educação superior integrante do sistema federal de ensino, de que trata o art. 2º do Decreto nº 9.235, de 15 de dezembro de 2017 (Brasil, 2020c).

Essa portaria, como observamos, autoriza o desenvolvimento de atividades que usem recursos tecnológicos, devido à impossibilidade da realização das aulas presenciais, mas não atribui uma denominação a tal tipo de ensino.

Em 18 de agosto de 2020, foi promulgada a Lei nº 14.040, “que estabelece normas educacionais excepcionais a serem adotadas durante o estado de calamidade pública” (Brasil, 2020d). No que se refere ao Ensino Superior, o artigo 3 menciona que não é necessário cumprir os dias de efetivo trabalho acadêmico, desde que a carga horária seja cumprida e que não haja prejuízo quanto aos conteúdos essenciais para o exercício da profissão (Brasil, 2020d).

Em 06 de outubro de 2020, o Conselho Nacional de Educação (CNE) emitiu o Parecer 15/2020, que, no capítulo 3, dedicado ao Ensino Superior, e Art. 26, determinou que “podem ser desenvolvidas atividades pedagógicas não presenciais vinculadas aos conteúdos curriculares de cada curso, por meio do uso de tecnologias da informação e comunicação, para fins de integralização da respectiva carga horária” (Brasil, 2020e, p. 14). Em todos os seus quatro artigos e nos 19 itens do Art. 3, a menção à atual modalidade de ensino se faz da seguinte forma: “aulas e atividades presenciais” que podem ser substituídas por “aulas e atividades não presenciais”; há também “atividades on-line síncronas” e “atividades online assíncronas”, “oferta

na modalidade a distância ou não presencial”; “aprendizado a distância ou não presencial”, “avaliações não presenciais”. Quanto aos recursos recomendados, há referência ao uso de “ambientes virtuais de aprendizagem”, e, em vários trechos, são mencionadas as “tecnologias digitais de informação e comunicação”, além do uso de “mídias digitais de longo alcance (*WhatsApp, Facebook, Instagram*)”, “laboratórios e equipamentos virtuais e tecnologias de interação”. Embora sejam esses os principais recursos, o material impresso também é indicado no texto, sobretudo, nas avaliações ou atividades de reforço. O regime domiciliar foi indicado para discentes positivos ou do grupo de risco e o teletrabalho para os profissionais do setor. A legislação não foi analisada exaustivamente, entretanto, diante dos documentos apresentados, é possível seguir a linha que o país adotou oficialmente. No que diz respeito ao Ensino Superior, é bastante evidente a recomendação do uso das tecnologias.

Refletindo sobre a situação em um contexto mundial, como se sabe, o vírus teve suas primeiras incidências na China, no final de 2019, foi se espalhando por diversos locais, tais como Europa e Estados Unidos, e mais tardiamente chegou ao Brasil, no início de 2020. Nesse sentido, ao passo que os países registravam casos, já procuravam combater o vírus e conter o seu avanço, bem como se reorganizavam internamente, considerando que a pandemia afetou os países em todos os seus setores. No âmbito educacional, identificou-se, já no início de 2020, o uso do termo em inglês *emergency remote teaching* (ERT), como o fizeram Hodges *et al.* (2020):

[...] *emergency remote teaching* (ERT) is a temporary shift of instructional delivery to an alternate delivery mode due to crisis circumstances. It involves the use of fully remote teaching solutions for instruction or education that would otherwise be delivered face-to-face or as blended or hybrid

courses and that will return to that format once the crisis or emergency has abated. The primary objective in these circumstances is not to re-create a robust educational ecosystem but rather to provide temporary access to instruction and instructional supports in a manner that is quick to set up and is reliably available during an emergency or crisis.

Embora o termo já tenha sido adotado na língua inglesa no início da pandemia, na legislação brasileira mencionada neste texto, ele não foi encontrado. Já nas publicações científicas, é possível localizá-lo já no primeiro semestre de 2020, como no ensaio de Arruda (2020), tratando do *ensino remoto emergencial*, tradução do termo em inglês “*emergency remote teaching*”. Em outros artigos, foi utilizada a sigla correspondente: ERE. O termo *ensino remoto* já era utilizado antes do período pandêmico, mas, nesse momento, tornou-se uma variante da nova forma *ensino remoto emergencial*. A forma expandida, com o acréscimo do adjetivo *emergencial*, reforça a ideia de que, após o término do período de crise, a situação se normalizará e as atividades retornarão ao modo presencial.

Ressaltamos que, nessa modalidade de ensino, não se faz uma mera transposição dos conteúdos para um suporte diferente do habitual, mas o ensino deve ser reestruturado em todos os seus aspectos. Assim, é importante distinguirmos Educação à Distância (EaD) do ensino online. No primeiro caso, o Decreto nº 9.057, de 25 de maio de 2017, define:

Art. 1º Para os fins deste Decreto, considera-se educação a distância a modalidade educacional na qual a mediação didático-pedagógica nos processos de ensino e aprendizagem ocorra com a utilização de meios e tecnologias de informação e comunicação, com pessoal qualificado, com políticas de acesso,

com acompanhamento e avaliação compatíveis, entre outros, e desenvolva atividades educativas por estudantes e profissionais da educação que estejam em lugares e tempos diversos (Brasil, 2017).

Em outras palavras, a EaD é “um tipo de ensino, com material pré-produzido, feito sem contato próximo entre professor e aluno” (Paiva, 2020, p. 60). No segundo caso, as denominações são várias, tais como “educação *on-line*, ensino mediado por computador, ensino *on-line*, aprendizagem *on-line*” (Paiva, 2020, p. 62).

A denominação mais recente – ensino remoto emergencial (ERE) – pode ser considerada uma forma de ensino online, porém, circunscrita ao período em que não houver um controle seguro da pandemia. Outra questão relevante é que a modalidade foi adotada sem preparação e programação prévia; os docentes e discentes, ao mesmo tempo em que estavam se adaptando a ela, participavam de sua construção, o que não era feito na EaD e no ensino online.

Como identificado na legislação brasileira, é indispensável o uso de recursos tecnológicos, mais especificamente das Tecnologias Digitais da Informação e Comunicação (TDICs), que compreendem

[...] uma gama de bases tecnológicas que possibilitam a partir de equipamentos, programas e das mídias, a associação de diversos ambientes e indivíduos numa rede, facilitando a comunicação entre seus integrantes, ampliando as ações e possibilidades já garantidas pelos meios tecnológicos (Soares *et al.*, 2015, p. 3).

Assim, a EaD vale-se do uso de tecnologias, principalmente dos ambientes virtuais de aprendizagem (AVAs), tais como o *Moodle*, *Google Sala de Aula*, entre outros (Paiva, 2020). E no caso do ensino online, é possível contar com

[...] ambientes virtuais onde são compartilhados vários tipos de recursos (vídeos, textos, imagens, animações, enquetes, avaliações, exercícios com resposta automática, plataforma para escrita colaborativa etc.) e oportunidades de interação (fórum, chat, e-mail) com colegas e professores de forma síncrona e assíncrona de qualquer lugar com acesso à internet (Paiva, 2020, p. 62).

O ERE acabou se aproveitando de todos os benefícios do que já vinha sendo feito anteriormente, e, conforme suas necessidades, foi impulsionando o aprimoramento das tecnologias, como o desenvolvimento de algumas extensões do *Google Meet*, por exemplo.

No caso específico da disciplina de *Lexicologia e Lexicografia*, a metodologia que utilizamos para conduzir as aulas não se restringiu apenas à exposição teórica, mas houve uma participação mais efetiva dos discentes, vindo ao encontro da seguinte reflexão:

Segundo BIGGS e TANG (2011), quanto mais ativa for a interação do aprendente com o processo de aprendizagem maior é a percentagem de aquisição de conhecimento. Um estudante mais bem aprende quando discute ativamente com outros um conceito do que quando ouve um professor numa sala de aula a falar sobre este conceito. O mesmo estudante compreende mais bem um conceito quando o aplica na prática, no seu dia a dia, do que se ler ou assistir a um vídeo (Casanova, 2020, p. 6).

Nessa perspectiva, propomos atividades de caráter mais prático, as quais detalhamos no próximo subitem. Essas adaptações foram motivadas pelo cenário que vivenciamos, o que fez com que nós, professores, em especial, saíssemos de nossa zona de conforto e assumíssemos novos papéis. O docente,

[...] mais do que transmitir conhecimentos, deve agora guiar o processo de aprendizagem do estudante de forma a desenvolver as suas capacidades, nomeadamente de aprender a aprender, da sua autoaprendizagem e da sua autonomia. O professor deve acompanhar, motivar, dialogar, ser líder e mediador, fomentando e mediando uma interação humana positiva (GOULÃO, 2012). Espera-se, ainda, que seja moderador, nas relações interpessoais e intrapessoais e faça o seu papel de auto e hetero-avaliador, de conteúdos e desempenhos. Espera-se também que sirva de suporte e estímulo aos estudantes, regulando e orientando as suas emoções, afetos e atitudes (DIAS, 2008). No professor recaem, pois, as funções de motivador, de criador de recursos digitais, de avaliador de aprendizagens e de dinamizador de grupos e interações online (Moreira; Henriques; Barros, 2020, p. 234).

Além de assumirmos um papel mais ativo, também devemos estimular nos discentes a também serem ativos e autônomos, o que faz com que o estudo deixe de ser uma atividade individual e para ser uma dinâmica e interativa.

Com base na legislação e nas considerações teóricas apresentadas, a próxima seção trata da realização da disciplina propriamente dita.

2. O CONTEXTO DA PÓS-GRADUAÇÃO NA UNIOESTE: A DISCIPLINA DE LEXICOLOGIA E LEXICOGRAFIA

Diante de todo esse percurso de reflexão sobre a nova realidade imposta pela pandemia e como o setor da Educação deveria proceder, a Unioeste, em relação às demais universidades do Estado, teve

uma decisão um pouco tardia quanto à retomada das aulas nas novas modalidades. Em 21 de maio de 2020, após ouvir professores e os representantes dos colegiados dos Programas de Pós-Graduação, a Resolução nº 52/2020 (UNIOESTE, 2020a) abriu a possibilidade de substituição das aulas presenciais por aulas remotas síncronas. A próxima etapa seria a ciência e a assinatura de termo de compromisso por parte dos discentes, afirmando possuírem recursos tecnológicos e demais condições para acompanhar as aulas. Após essa resolução, em 24 de junho de 2020, o Programa de Pós-Graduação em Letras, *campus* de Cascavel (PR), lançou o Edital 22/2020 (UNIOESTE, 2020b), orientando e normatizando as inscrições nos componentes curriculares do novo período de atividades. Como ressaltando, a disciplina de *Lexicologia e Lexicografia* foi realizada no período de 07 de julho a 01 de setembro de 2020, contando com 21 alunos concluintes. A seguir, refletimos sobre três momentos das atividades: (i) a preparação; (ii) o desenvolvimento; (iii) a avaliação da disciplina.

2.1. A preparação

A proposta da disciplina foi adequada ao novo formato, sendo constituída por aulas remotas síncronas e por atividades assíncronas, a fim de completar a carga horária total. Em um total de nove encontros, sete foram dedicados às aulas expositivas dialogadas e dois foram dedicados aos seminários em grupo. Para a realização das atividades, selecionamos duas ferramentas: o *Google Sala de Aula*, em que eram colocados avisos, informações sobre a disciplina e as atividades avaliativas propostas, e o *Google Meet*, utilizado para as aulas remotas síncronas.

As sete primeiras aulas foram divididas em três momentos: (i) atividade prévia, com base na leitura dos textos teóricos, disponibili-

zados no *Google Sala de Aula*; (ii) aula remota síncrona, com 1h30min de duração, utilizando-se o *Google Meet*; (iii) atividade prática após a aula, por meio do *Google Sala de Aula*.

Os conteúdos explorados nas sete primeiras aulas foram:

- 1) introdução às ciências do léxico;
- 2) história do português brasileiro;
- 3) a unidade lexical e as relações de significação;
- 4) renovação do léxico;
- 5) lexicografia;
- 6) arquitetura das obras lexicográficas;
- 7) lexicografia pedagógica.

Selecionamos esses temas porque, na graduação em Letras, oferecida na mesma instituição e *campus*, não há disciplinas que abordem tais estudos. Assim, julgamos necessário apresentar, além de noções basilares, um panorama dos estudos que são realizados nos âmbitos nacional e internacional.

Para um bom andamento da disciplina, tendo em vista a nova realidade, os discentes contaram com os seguintes instrumentos norteadores:

- a. programa da disciplina, com datas, temas e textos que deveriam ser lidos para acompanhar cada encontro;
- b. diretrizes sobre o funcionamento da disciplina, divididas nas seguintes seções: 1) Sobre plano de ensino: 1a) Metodologia; 1b) Avaliação; 1c) Frequência; 2) Sobre Plataformas: 2a) Sobre *Google Sala de Aula*; 2c) Sobre *Google Meet*;

c. diretrizes e critérios de avaliação para os seminários em grupo.

Em um primeiro momento, tais documentos foram enviados por e-mail e posteriormente postados no *Google Sala de Aula*. Além disso, por meio de dois instrumentos elaborados no *Google Formulários*, obtivemos alguns dados dos cursistas. O primeiro, *Dados iniciais dos discentes*, solicitava informações acerca da formação e atuação profissional com o ensino de línguas. O segundo – *Questionário sobre a disciplina de Lexicologia e Lexicografia (Unioeste, modalidade remota, 2020)* – foi mais completo, em que os discentes foram convidados a avaliar alguns aspectos da disciplina. Exploramos esse documento e os respectivos dados na seção 2.3.

Todas essas ferramentas foram elaboradas com o objetivo de guiar os discentes no acompanhamento da disciplina, já que, no ensino presencial, muitas informações são fornecidas e dúvidas são esclarecidas nesse contato. Além disso, possibilitaram que otimizássemos o tempo. Como as aulas remotas síncronas foram reduzidas em relação às aulas tradicionais, tendo todas as questões de ordem burocrática resolvidas, o tempo em que todos estão juntos é melhor aproveitado. Tal organização vai ao encontro do que defendem Moreira, Henrique e Barros (2020, p. 354):

Estas indicações são fundamentais para desenvolver processos de autorregulação e de autonomia nos estudantes e para uma navegação sem sobressaltos no ambiente digital. Para reforçar esta orientação na navegação é também útil criar um espaço de comunicação assíncrona, um quadro de avisos e notícias no tópico inicial do ambiente, onde o professor possa colocar informação acerca de uma determinada atividade que irá iniciar.

Os recursos utilizados para a comunicação e orientação foram o mural do *Google Sala de Aula* e o e-mail.

Após descrevermos a preparação da disciplina e das aulas, a seguir, relatamos o seu desenvolvimento.

2.2. O desenvolvimento

Conforme mencionamos na seção 2.1, as sete primeiras aulas foram divididas em três momentos. No primeiro, era obrigatória a leitura dos textos de apoio para que fosse possível acompanhar e participar das discussões nas aulas remotas síncronas. A fim de garantir tal tarefa, foi elaborada uma atividade, uma pergunta norteadora posta no *Google Sala de Aula*. Os alunos deveriam percorrer um pouco sobre os textos, apresentando também uma visão crítica sobre eles.

Como exemplo, na aula 2, baseados nos textos de Biderman (2002) e Murakawa (2015), foi feita a seguinte proposta, que deveria ser respondida no limite de 15 a 20 linhas:

De acordo com os textos da aula 2, responda: 1) Quais os aspectos importantes com relação à formação e consolidação da norma lexical lexicográfica no Português do Brasil? 2) Que aspectos chamam mais atenção na elaboração do Dicionário Histórico de Biderman/Murakawa?

Nas respostas, houve bastante menção ao percurso feito pela língua falada no país, até seu estabelecimento como norma nacional, inclusive marcada pela elaboração e pela publicação de diversas obras lexicográficas. Já no segundo texto, os aspectos ressaltados dizem respeito especialmente ao banco de dados criado para elaboração da obra. No entanto, notamos certa dificuldade em tratar de detalhes da sua elaboração, já que a maioria dos discentes não estava familiarizada

com esse aspecto, como a questão dos critérios para se determinar a homonímia e as formas variantes, bem como explicações a respeito da datação e de outros aspectos (Murakawa, 2015).

A atividade era individual e não havia acesso às respostas pelo grupo. Outra exigência era a entrega do texto antes do segundo momento, que foi a aula remota síncrona, realizada pelo *Google Meet*, com duração de 1h30min. Nesse momento, realizamos a exposição teórica com base nos textos indicados, bem como em outros relacionados ao tema, além de exemplos práticos, sempre que possível. Para auxiliar a explanação, elaboramos slides, que tornaram mais claras as explicações e foram um ponto de apoio para os discentes. Nesse contexto, muitos, além de se atentarem para a exposição, tiravam fotos da tela, a fim de economizar o tempo de cópia. Após essa parte, havia espaço para comentários e perguntas. Em 2020, ao utilizarmos a ferramenta do *Google Meet*, enquanto se compartilhava uma apresentação, não era possível ver os participantes, nem mesmo o *chat* do programa. Então, algumas perguntas que eram feitas durante a exposição acabavam interferindo um pouco na gestão do tempo. Naquele momento, não utilizamos duas telas, por exemplo, a do computador e a celular para solucionar o problema. Em 2021, contudo, apareceram outras soluções.

Na aula sobre renovação do léxico, alguns exemplos foram apresentados. Um termo bastante recente seria o de “covidiotas”, encontrado na internet. Origina-se da língua inglesa, “*covidiot*”, formado por dois substantivos “*covid*” (relativo ao vírus da atual pandemia) e “*idiot*”. Em 2020, a palavra era uma novidade, mas, em 2021, já se encontra dicionarizada na língua inglesa, em suporte online, com a seguinte acepção: “*a person who annoys other people by refusing to obey the social distancing rules designed to prevent the spread of COVID-19*” (Covidiot, 2021). No dicionário, há a referência ao ano de surgimento da palavra, que seria o ano de 2020, bem como há a indicação das

marcas de uso “*informal*” (informal) e “*disapproving*” (inadequado). Buscas na rede levaram à conclusão de que a palavra ainda não foi incorporada por suas obras lexicográficas.

No terceiro momento, propusemos uma atividade de caráter mais prático, não muito comum em cursos de Pós-Graduação em Letras. Geralmente, são solicitados textos acadêmicos, como artigos finais relacionados à disciplina ou, se possível, à pesquisa do discente. Este é um exemplo do que foi solicitado:

1) Escolha uma palavra da língua portuguesa de origem indígena ou africana. 2) Consulte tal palavra em dois dicionários da Língua Portuguesa disponíveis na internet: obrigatoriamente o “Vocabulário Portuguez e Latino”, de Raphael Bluteau (1728) e um atual (Aurélio, Priberam, Michaelis). 3) Copie as acepções. 4) Faça um comparativo das acepções da palavra selecionada nas duas obras. Não ultrapasse uma página. Obs.: O dicionário de Bluteau está no link <https://www.bbm.usp.br/pt-br/dicionarios/>.

Para realizar essa atividade, vários entendimentos foram necessários. O primeiro deles era encontrar palavras do português do Brasil, de origem indígena ou africana, desconsiderando nomes de pessoas ou lugares, já que a disciplina não tratou da Onomástica. Sabe-se que essa informação pode ser facilmente encontrada na internet. No entanto, é necessário confirmá-la, consultando a etimologia das palavras em dicionários confiáveis. Selecionada a palavra, era necessário copiar as acepções do dicionário de Bluteau (1712-1728), disponível na página da Biblioteca Brasileira Guita e José Mindlin. No site, é possível fazer a busca pelos verbetes, digitando o que se procura. Alguns dos itens lexicais pesquisados foram: “mandioca”, “moleque”, entre outros. Um exemplo interessante foi o de “cuya”. Na obra mais antiga, encontra-se:

Figura 1 – Verbete “Cuya”

CUYA, Cùya. Vaso de barro, em que bebe o Gentio do Brasil. Rede, cabaço, & Cuya. Vasconc. Notic. do Brasil pag. 123.

Fonte: Bluteau (1712-1728, p. 648)

O Dicionário Aurélio online apresenta duas acepções:

A casca do fruto da cuieira, que, depois de seca e limpa, é empregada pelos indígenas como prato etc.; cuité.

Vasilha feita da cabaça, usada em vários misteres, e utilizada no Rio Grande do Sul, ricamente adornada, para preparar e beber o chimarrão.

[Brasil] Gír. Cabeça, coco.

A primeira diferença notável é a ortografia. A primeira obra indica apenas uma acepção e, em seguida, uma abonação, em que é mencionada também a fonte. Na segunda obra, são duas acepções indicadas pelo fim de um parágrafo e início do outro. A primeira acepção coincide nas duas obras, mas as definições são diferentes. O Dicionário Aurélio acrescenta mais pormenores, como o material com que é feita a cuia, geralmente utilizada por indígenas ou pelo “Gentio do Brasil”, como diz Bluteau (1712-1728). A segunda mostra que a palavra manteve o primeiro sentido e adquiriu outro, também sendo um utensílio, feito de outro material, utilizado por várias profissões e também no Rio Grande do Sul, para o preparo e consumo do chimarrão.

É claro que a escolha das palavras e das obras de consulta ficou a critério de cada discente, o que proporcionou diversidade e riqueza

za de resultados. Além disso, embora algumas palavras tenham sido tema de mais de uma atividade, como “moleque”, as reflexões foram distintas. Certa dificuldade foi demonstrada por alguns, ao realizarem a busca no arquivo do dicionário Bluteau, o que foi sanado com mais explicações individuais dadas por meio do *Google Sala de Aula*.

No conjunto de todas as atividades práticas solicitadas, essa foi a que se destacou com os melhores resultados. O empenho na realização resultou na compreensão do tema, ficando evidente que cada palavra tem sua história. Todo esse empenho por parte dos discentes necessitou de uma contrapartida, a fim de garantir o acompanhamento e a compreensão dos temas durante a realização da disciplina. Nesse sentido, para cada atividade realizada (prévia ou após as aulas), houve um retorno, de nossa parte, com comentários, sugestões ou esclarecimentos a respeito dos temas, principalmente pelo *Google Sala de Aula*. Isso exigiu também intenso trabalho, recompensado pelos resultados.

Todas as atividades apresentadas compuseram o conjunto de avaliações da disciplina. Para concluí-la, as duas últimas aulas foram dedicadas a seminários em grupo. Os discentes deveriam analisar, preferencialmente, dicionários de língua geral online e apresentar suas reflexões à turma, em uma exposição de 15 a 20 minutos. Cada grupo deveria basear-se no seguinte roteiro:

O seminário deverá discutir as seguintes questões e, dar exemplos, quando necessário:

- a) tipologia do dicionário, a quem se destina;
- b) organização geral do dicionário (todas as informações e anexos que o dicionário traz);
- c) macroestrutura;

- d) microestrutura;
- e) sistema de remissivas;
- f) análise de alguns verbetes.

Nesse caso, a avaliação foi dividida igualmente em duas partes: a primeira, com base no PDF dos slides, correspondeu ao trabalho em grupo; a segunda, com base na apresentação individual.

Os dicionários escolhidos pelos alunos foram: *Wordreference*, na seção do dicionário monolíngue de inglês; Priberam; Dicionário Aurélio (5ª edição, via aplicativo para celular); *Merriam Webster Visual Dictionary Online* (2ª edição); Dicionário Brasileiro da Língua Portuguesa (Michaelis); *Diccionario de la Real Academia Española* (23ª edição). Nota-se que, entre as seis obras escolhidas, três tratavam da língua portuguesa, uma da língua espanhola e duas da língua inglesa. Com relação ao dicionário visual, como o próprio nome diz, seu ponto de partida são as imagens, apoiando-se na multimodalidade. Embora tal questão teórica não tenha sido abordada na disciplina, o grupo se empenhou e apresentou algumas reflexões a esse respeito.

Além dessas atividades específicas, o processo avaliativo ocorreu de forma constante, pois contou com atividade baseada nos textos teóricos, com atividade prática após a aula remota síncrona, ambas individuais, e finalmente com o seminário em grupo, avaliado no âmbito do grupo e também de forma individual.

A seguir, tecemos alguns comentários sobre a avaliação da disciplina.

2.3. A avaliação da disciplina

A fim de ter uma visão, por parte dos discentes, sobre a eficácia da realização da disciplina neste novo formato, utilizamos um formulário do *Google*. Ele foi composto por 25 questões, divididas em três sessões: a) Geral (a fim de obter dados pessoais dos discentes); b) Quanto aos aspectos práticos; c) Quanto ao conteúdo e às suas aplicações. Algumas questões foram de múltipla escolha e outras abertas. Embora não fosse obrigatório respondê-lo, houve 21 respostas. Apresentamos, a seguir, alguns pontos relevantes.

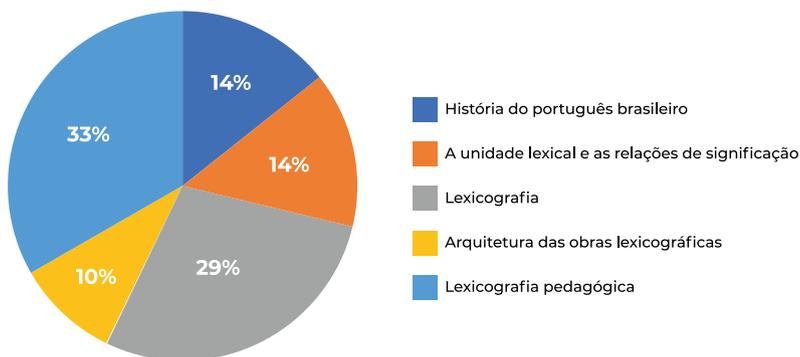
No item Geral, o dado mais notório é que, embora a disciplina tenha sido realizada pelo formato remoto, a maioria dos inscritos era originária de cidades da região, havendo dois deles que eram da cidade de Dourados, Mato Grosso do Sul. No entanto, com relação aos anos anteriores, foram mais inscrições, tanto de alunos regulares quanto de especiais. Certamente, se houvesse uma maior divulgação, o número de inscritos seria maior.

No item relativo aos aspectos práticos, o uso do *Google Sala de Aula* foi aprovado por todos, pelo aspecto da comunicação e pela possibilidade do retorno sobre as atividades por parte da docente. O sistema avaliativo foi igualmente aprovado, enfatizando-se o fato de todos terem tido de se dedicar intensamente ao cumprimento de todas as tarefas. Entretanto, como se sabe, o ensino remoto não é perfeito. Foram relatadas dificuldades quanto à conexão individual da internet (28,6% das respostas), bem como com dificuldades com o uso dos aparelhos eletrônicos, sejam computadores ou celulares, com a mesma porcentagem. Outro comentário encontrado em um formulário diz respeito ao tempo destinado às dúvidas e aos comentários durante a aula remota síncrona. Nesse caso, o discente disse que esse espaço

deveria ser feito durante a exposição teórica e não assim que ela fosse concluída, em cada encontro.

Com relação ao conteúdo, aos objetivos da disciplina, ao programa, à distribuição do conteúdo no decorrer da disciplina e aos textos indicados para leitura, a aprovação foi total. Já os temas preferidos, foco da questão 22, são indicados no gráfico a seguir:

Gráfico 1 – Temas preferidos pelos discentes



Fonte: Elaborada pela autora (2021), com auxílio do *Google Formulários*

Os temas “Introdução às ciências do léxico” e “Renovação do léxico” não foram mencionados.

No que tange às contribuições da disciplina, 52,4% afirmaram que essas se dão no ensino em sala de aula, 33,3% apontaram a pesquisa na Pós-Graduação, 9,6% indicaram ambos os âmbitos e 4,8% as relacionaram aos estudos da tradução.

Por fim, na questão sobre qual aspecto foi mais relevante na disciplina, destacamos dois comentários: “Alguns dos textos eram bastante densos. Realizar a atividade anteriormente à aula facilitava um pouco a compreensão textual e as explicações sanavam as dúvidas. Achei muito interessante a dinâmica dos seminários, ocasião em que

a teoria apresentou melhor sentido” (DISCENTE A). Nesse caso, são ressaltadas as atividades anteriores à aula e os seminários. Já o outro comentário mostra a preferência pelas atividades realizadas após as aulas expositivas: “O que mais me chamou a atenção e eu gostei, foram as atividades práticas no final de cada aula. Assim, após ler, comentar e discutir a teoria, foi possível fazer breves análises colocando em prática o conteúdo visto” (DISCENTE B). De modo geral, o que se notou foi a aceitação geral do conjunto das atividades propostas.

CONSIDERAÇÕES FINAIS

Considerando que essa foi nossa primeira experiência de aulas remotas na Pós-Graduação, é possível afirmar que os resultados foram bastante satisfatórios. A partir da data em que foi aberta a possibilidade das disciplinas de Pós-Graduação na Unioeste em formato remoto, houve pouco tempo para uma reflexão aprofundada sobre como proceder, sem modelos dos quais partir. Embora o formato proposto tenha tido grande aceitação, demandou grande dedicação, sobretudo, de tempo de ambas as partes, docente e discentes, o que não seria possível, caso não houvesse esse espaço no cotidiano de cada um dos envolvidos.

De acordo com relatos da maioria dos discentes, as atividades mais práticas foram as que mais despertaram interesse. Com relação à primeira, baseada na leitura dos textos teóricos, seria uma atividade mais tradicional, já que se parte da premissa que os textos serão lidos para que se possa seguir as discussões das aulas. Em se tratando das atividades práticas propostas após as aulas, bem como no caso dos seminários, buscar e discutir os exemplos sempre requereram a utilização da teoria, o que possibilitou uma melhor compreensão do conteúdo.

Houve três momentos de interação: durante as aulas remotas síncronas, entre docente e discentes; no esclarecimento de dúvidas e retorno sobre as atividades no *Google Sala de Aula*; e nos seminários em grupo, em que os discentes tiveram de se organizar e formular uma opinião conjunta sobre o material analisado.

Embora essa oferta tenha sido realizada em um contexto de pandemia, fica evidenciada a democratização do ensino por meio da internet. No contexto das aulas presenciais, há a questão do deslocamento, o que por vezes requer dispensa de atividade de trabalho e principalmente uso de recursos financeiros. Igualmente foi uma possibilidade de difundir os estudos da área no país, até mesmo pelo fato de a divulgação na rede ter maior alcance. Por fim, na esperança de que o ensino retorne à normalidade, sugerimos que essa porta da modalidade remota permaneça aberta.

REFERÊNCIAS

ARRUDA, Eucídio Pimenta. Educação remota emergencial: elementos para políticas públicas na educação brasileira em tempos de Covid-19. **EmRede - Revista de Educação a Distância**, v. 7, n. 1, p. 257-275, jan./jun. 2020.

BIDERMAN, Maria Tereza Camargo. A formação e a consolidação da norma lexical e lexicográfica no português do Brasil. In: NUNES, J. H.; PETTER, M. (Orgs.). **História do saber lexical e constituição de um léxico brasileiro**. São Paulo: Humanitas/FFLCH/USP/Pontes, 2002. p. 65-82.

BLUTEAU, Rafael. **Vocabulário Português Latino**. Lisboa: Oficina de Pascoal da Sylva, 1712-1728. Disponível em: <https://www.bbm.usp.br/pt-br/dicionarios/vocabulario-portuguez-latino-aulico-anatomico-architectonico/>. Acesso em: 25 fev. 2021.

BRASIL. Presidência da República. **Decreto nº 9.057, de 25 de maio de 2017**. Regulamenta o art. 80 da Lei nº 9.394, de 20 de dezembro de 1996, que estabelece as diretrizes e bases da educação nacional. Brasília: Presidência da

República, 2017. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2017/decreto/d9057.htm. Acesso em: 01 out. 2021.

BRASIL. Ministério da Educação. **Portaria nº 343, de 17 de março de 2020**. Dispõe sobre a substituição das aulas presenciais por aulas em meios digitais enquanto durar a situação de pandemia do Novo Coronavírus - COVID-19. Brasília: MEC, 2020a. Disponível em: <https://www.in.gov.br/en/web/dou/-/portaria-n-343-de-17-de-marco-de-2020-248564376>. Acesso em: 25 fev. 2021.

BRASIL. Congresso Nacional. **Decreto Legislativo nº 6, de 2020**. Reconhece, para os fins do art. 65 da Lei Complementar nº 101, de 4 de maio de 2000, a ocorrência do estado de calamidade pública, nos termos da solicitação do Presidente da República encaminhada por meio da Mensagem nº 93, de 18 de março de 2020. Brasília: Congresso Nacional, 2020b. Disponível em: http://www.planalto.gov.br/ccivil_03/Portaria/DLG6-2020.htm. Acesso em: 25 fev. 2021.

BRASIL. Presidência da República. **Portaria nº 544, de 16 de junho de 2020**. Reconhece, para os fins do art. 65 da Lei Complementar nº 101, de 4 de maio de 2000, a ocorrência do estado de calamidade pública, nos termos da solicitação do Presidente da República encaminhada por meio da Mensagem nº 93, de 18 de março de 2020. Brasília: Presidência da República, 2020c. Disponível em: <https://www.in.gov.br/en/web/dou/-/portaria-n-544-de-16-de-junho-de-2020-261924872>. Acesso em: 25 fev. 2021.

BRASIL. Senado Federal. **Lei nº 14.040, de 18 de agosto de 2020**. Estabelece normas educacionais excepcionais a serem adotadas durante o estado de calamidade pública reconhecido pelo Decreto Legislativo nº 6, de 20 de março de 2020; e altera a Lei nº 11.947, de 16 de junho de 2009. Brasília: Senado Federal, 2020d. Disponível em: <https://legis.senado.leg.br/norma/32560565>. Acesso em: 30 set. 2021.

BRASIL. Conselho Nacional de Educação. **CNE/CP nº 15/2020, de 06 de outubro de 2020**. Reorganização do Calendário Escolar e da possibilidade de cômputo de atividades não presenciais para fins de cumprimento da carga horária mínima anual, em razão da Pandemia da COVID-19. Brasília: CNE, 2020e. Disponível em: http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=160391-pcp015-20&category_slug=outubro-2020-pdf&Itemid=30192. Acesso em: 30 set. 2021.

CASANOVA, Diogo. Um modelo de design educacional para ambientes mistos e EaD. **EmRede - Revista de Educação a Distância**, v. 7, n. 1, p. 4-20, jan./jun. 2020.

COVIDIOT. *In*: The Oxford Learner's Dictionaries. Oxford University Press. Disponível em: <https://www.oxfordlearnersdictionaries.com/definition/english/covidiot?q=COVIDIOT>. Acesso em: 31 ago. 2023.

HODGES, Charle; MOORE, Stephanie; LOCKEE, Barb; TRUST, Torrey; BOND, Aaron. The Difference Between Emergency Remote Teaching and Online Learning. **Educause Review**, mar. 2020. Disponível em: <https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning>. Acesso em: 25 fev. 2021.

MOREIRA, Jose Antônio Marques; HENRIQUES, Susana; BARROS, Daniela. Transitando de um ensino remoto emergencial para uma educação digital em rede, em tempos de pandemia. **Dialogia**, São Paulo, n. 34, p. 351-364, jan./abr. 2020.

MURAKAWA, Clotilde. Dicionário Histórico do Português do Brasil: testemunho lexical da língua portuguesa no Brasil Colônia. **Debate Terminológico**, n. 14, p. 75-88, 2015.

PAIVA, Vera Lúcia Menezes de Oliveira e. Ensino remoto ou ensino a distância. Efeitos da pandemia. **Estudos Universitários** - Revista de Cultura, v. 37, n. 1-2, p. 58-70, dez. 2020.

SOARES, Simária de Jesus *et al.* O uso das tecnologias digitais de informação e comunicação no processo de ensino-aprendizagem. *In*: CONGRESSO INTERNACIONAL ABED – EDUCAÇÃO A DISTÂNCIA, 21., 2015. **Anais** [...]. Bento Gonçalves: ABED, 2015. Disponível em: http://www.abed.org.br/congresso2015/anais/pdf/BD_145.pdf. Acesso em: 25 fev. 2021.

UNIOESTE. Universidade Estadual do Oeste do Paraná. **Resolução nº 52/2020-CEPE, de 21 de maio de 2020**. Regulamenta a possibilidade de substituição de aulas presenciais por aulas remotas síncronas, em caráter excepcional, nos programas e nos cursos de pós-graduação stricto e lato sensu da Unioeste, durante a suspensão das atividades acadêmicas letivas presenciais, determinadas pelo Ato Executivo n 21/2020-GRE. Cascavel: Unioeste, 2020a. Disponível em: <https://midas.unioeste.br/sgav/grupo#/busca/?grupo=ResolucaoCEPE&unidade=Reitoria&ano=2020&tipoarquivo=18>. Acesso em: 09 out. 2020.

UNIOESTE. Universidade Estadual do Oeste do Paraná. Programa de Pós-Graduação em Letras. **Edital 022/2020-PPGL, de 24 de junho de 2020**. Súmula: Convocação de Alunos Regulares para Inscrição em Disciplinas Obrigatórias, Eletivas e Seminários Avançados para o ano de 2020 – Aulas

Remotas Síncronas do Programa de Pós-Graduação em Letras, Área de Concentração em Linguagem e Sociedade, nível de mestrado e doutorado. Cascavel: Unioeste/PPGL, 2020b. Disponível em: https://www5.unioeste.br/portalunioeste/arq/files/PPGL/Edital_022-2020-PPGL_-convocacao_alunos_regulares_matriculas_disciplinas_remoto_sincrono_2020.pdf. Acesso em: 01 out. 2021.

(1980) [1585], I – DO CLIMA E TERRA DO BRASIL - E DE ALGUMAS COUSAS NOTÁVEIS QUE SE ACHÃO ASSI NA TERRA COMO O MAR [A00_0749 p. 59].

PÃO-DE-AÇÚCAR

Açúcar resfriado no tendal das primitivas casas de engenho, depois de clareado na casa de purgar, era posto em longas formas dentro das quais tomava consistência resultando um bolo, donde o seu nome. [...] que o lavrador será obrigado a plantallas de cannas que não poderá moer mais do que no Engenho do proprietario que pelas moer lhe pertence a metade do assucar que produzirão, além do que lhe ha da de dar mais daquella metade e com que ficou hum **pão de assucar** por cada quinze: e isto pela renda da terra [...]. LUIZ DOS SANTOS VILHENA (1921) [1802], CARTA QUINTA [A00_0407 p. 182]. 1ª. datação [1552] Ó trato bendito, não de açuquere corruptivel, mas de graça, mais saborosa que favo de mel, quam poucos há que te queirão ter! P. MANUEL DA NÓBREGA (1956) [1552], CARTA DO P. MANUEL DA

PARTE II LEXICOGRAFIA

OBRAS DE REFERÊNCIA PORTUGUESAS NO *HOBSON-JOBSON*: UMA ABORDAGEM LEXICOGRÁFICA

Maria Celeste Augusto

“[Colóquios] A most valuable book, full
of curious matter and good sense”
(Yule, 1886, p. XXXV)

INTRODUÇÃO

Certos textos de autores portugueses tiveram uma grande relevância na elaboração das entradas do Glossário de Indo-Inglês *Hobson-Jobson*. Abordar todas as obras e autores portugueses citados no referido glossário é, como dizem os franceses, “la mer à boire”. Por isso, decidiu-se trabalhar apenas os dois títulos portugueses mais mencionados no *Glossary: Colóquios dos Simples e Drogas* de Garcia de Orta e *Décadas da Asia* de João de Barros. Pretendemos primordialmente refletir sobre o modo como os autores do *Glossary* se serviram dos referidos textos portugueses. Como segundo propósito, procurar-se-á justificar o epíteto *alambique* dado pelos dois lexicógrafos ingleses à língua portuguesa, mais precisamente ao Indo-Português. É a função de veículo transmissor do Português, ao servir de intermediário entre as línguas “exóticas” e as línguas europeias, que não tem recebido, salvo raríssimas exceções, a atenção que merece na maior parte dos estudos linguísticos, sobre essa matéria.

Após este preâmbulo, no parágrafo seguinte, será caracterizado o *Glossary* e dada atenção às fontes, sobretudo portuguesas, usadas por Henry Yule e A.C. Burnell; no momento seguinte, far-se-á uma reflexão sobre a função e o modo como foi feito o aproveitamento lexicográfico das obras em questão. Segue-se uma fundamentação para o qualificativo *alambique* atribuído pelos autores do *Glossary* ao português. O trabalho termina com uma síntese conclusiva.

1. CARACTERIZAÇÃO DO *HOBSON-JOBSON*

1.1. Título e edições

Hobson-Jobson é o título alternativo de um glossário de Indo-Inglês, cujo nome completo é *A Glossary of Colloquial Anglo-Indian Words and Phrases, and of Kindred Terms, Etymological, Historical, Geographical and Discursive*. Há, porém, um terceiro título de extensão intermédia, que aparece registado na primeira página do *Glossary*, propriamente dito, e que é o seguinte: *A Glossary of Anglo-Indian Colloquial Terms and Phrases of Analogous Origin*. Esse e o título mais longo elucidam de antemão o leitor acerca do conteúdo da obra. O mesmo não acontece com o título alternativo *Hobson-Jobson*, que é, no entanto, o mais conhecido. Na entrada respectiva o segmento denominativo é considerado como “one of the most highly assimilated class of Anglo-Indian argot, [...]. It is peculiar to the British soldier and his surroundings”. Esse processo de adaptação fonológica e ortográfica à língua importadora de um segmento de uma outra língua encontra-se profusamente ilustrado no *Glossary*, e está registado no *Oxford English Dictionary* (OED) como *The law of Hobson-Jobson* ou *Hobson-Jobsonism*.

O *Glossary* teve a sua primeira edição em 1886, em vida de apenas um dos seus autores, Henry Yule. A. Burnell falecera em 1882. A base linguística de todo o *Glossary* é da autoria de A.C. Burnell, ficando o trabalho final todo a cargo de Yule. A obra apresenta algumas observações introdutórias, que são seguidas por dois tipos de “Notes”: A e B. A lista de obras citadas inclui vários relatos de viagens, inúmeros dicionários como *Vocabulario Portuguez Latino* (Bluteau, 1712-1728), *Dictionnaire de la Langue Française* (Littré, 1873-1877) e *Etymologisches Wörterbuch der Romanischen Sprachen* (Diez, 1869). Quanto à datação das citações verifica-se alguma inconsistência, aliás, previamente indicada pelos autores, visto que, por vezes, a data indicada é a do uso da palavra ou expressão e não a da publicação da obra, um aspecto que gera alguma confusão. Todavia, na lista bibliográfica é dada sempre a data da publicação. Ao contrário da definição usual de glossário enquanto ‘vocabulário explicativo anexo a um texto’ o *Hobson-Jobson*, embora chamado *Glossary*, não é um glossário vulgar não se reportando a uma obra única mas a um conjunto imenso de textos versando todo o tipo de matérias, cujo denominador comum é a sua relação com o Oriente. Também não é uma enciclopédia nem um dicionário etimológico, contudo, o espírito enciclopédico com que os autores percorreram toda uma série de obras de referência e o registo abrangente de várias hipóteses etimológicas e de múltiplos extractos ilustrativos, revelam-se ao longo das 1000 páginas e cerca de 2000 entradas do *Glossary* e fazem dele um repositório único de vários saberes.

Uma segunda edição, melhorada, foi feita em 1903 pelo orientalista William Crooke, que acrescentou um prefácio. Crooke não alterou nem a macroestrutura nem a microestrutura, limitando-se a fazer correções, sobretudo nas datas e nas citações, porém, nada omitindo. Todos os acrescentos e melhoramentos feitos por Crooke estão

registados entre parêntesis rectos e consistem em: verificação de citações da primeira edição, registo de novas citações, inclusão de alguns verbetes acidentalmente omitidos e de algumas novas referências. Também introduziu um índice (Crooke, 1903, p. 987-1021), feito por Charles Partridge, de todos os termos e nomes, analisados ou não, incluídos nas microestruturas. Essa edição, para além de dois prefácios, um de Yule e outro de W. Crooke, reproduz os *Introductory Remarks*, da edição de 1886 (Yule, 1886, p. xv-xxv) que incluem a **Nota A** e a **Nota B**, acima referidas. A Nota B sobre o Indo-Português foi feita por A. C. Burnell. Diga-se que a base linguística de todo o *Glossary* é da autoria deste último, que morreu quatro anos antes da sua publicação, ficando o trabalho final todo a cargo de Yule. Às *Introductory Remarks* segue-se uma lista completa dos títulos dos livros citados no *Glossary* e finalmente há uma corrigenda. O *Glossary* propriamente dito apresenta 987 páginas a duas colunas e ainda o índice feito por C. Partridge. A versão em linha feita pela Universidade de Chicago facilitou a contagem dos dados referentes quer ao *Hobson-Jobson* quer às obras e autores portugueses selecionados para este trabalho. H. Yule colaborou ainda na 9ª edição da *Enciclopédia Britânica*, com uma série de artigos relativos, nomeadamente, à Ásia e a Viajantes, (Teltscher, 2011). Pelo facto de conhecer bem James Murry, editor do NED, mais tarde OED, cuja produção coincidiu em parte com a do *Glossary*, Yule enviou-lhe algumas entradas deste último a fim de ter o parecer daquele editor. Essas entradas foram introduzidas no OED por Murry. Dessa situação resultou uma intensa colaboração entre os dois lexicólogos/lexicógrafos e a inclusão no OED de uma boa quantidade de entradas, originariamente pertencentes ao *Glossary*. São para cima de 300 entradas reconhecidas, porém, muitas outras há oriundas do *Glossary*, cuja origem não está devidamente assinalada.

Recentemente, em 2013, já depois de se ter feito uma versão simplificada, mas não publicada, apresentada num Congresso em 2011, foi publicada uma nova edição do *Glossary*; edição bastante reduzida (quase metade da edição de 1903), preparada por Kate Teltscher. Além de manter os dois prefácios da edição de 1903 e a Introdução feita por Yule em 1886, é acompanhada por outra introdução feita pela nova editora, que versa não só sobre os dois autores Yule e Burnell, como sobre os méritos, falhas e omissões da obra. Teltscher (2013) refere-se ainda à recepção que a obra teve, às inúmeras citações que atestaram a vastidão de conhecimentos dos seus autores e ainda às suas atitudes coloniais e a certa arrogância cultural, que mais não é do que uma característica, fruto do discurso imperial da época. Assim, por exemplo, na entrada da palavra *opium*, matéria importante na época, os autores decidiram omitir o facto de a Inglaterra ter o monopólio da sua exploração.

Uma vez que a obra foi reduzida a cerca de um terço houve naturalmente uma seleção de entradas. Essa seleção fez-se, segundo a própria editora, com base na sua importância histórica, literária e linguística. Os aspectos negativos do *Glossary*, porém, não impediram o linguista Braj B. Kachru de, em 1980, considerar o *Hobson-Jobson* “the only existing dictionary which closely approximates serious lexical work in Indian English” (Teltscher, 2013, p. xxi). O facto de a edição de Teltscher em 2013 não estar digitalizada e de fugir em grande parte à edição completa de 1903, especialmente quanto ao número de entradas, pois como acima se disse está reduzida a um terço, levou-nos a manter, como base deste estudo, essa edição que, embora muito melhorada por W. Crooke, é a que segue mais de perto o espírito da edição feita por Yule, em 1886.

1.2. Conteúdo

Ao contrário da definição usual de glossário enquanto ‘vocabulário explicativo anexo a um texto’, o *Hobson-Jobson*, embora chamado *Glossary*, não é um glossário vulgar não se reportando a uma obra única, mas a um conjunto imenso de textos versando todo o tipo de matérias, cujo denominador comum é a sua relação com o Oriente. Embora apresente uma estrutura feita de entradas ordenadas alfabeticamente, também não é uma enciclopédia nem um dicionário etimológico, contudo, o espírito enciclopédico com que os autores percorreram toda uma série de obras de referência e o registo abrangente de várias hipóteses etimológicas e de múltiplos extratos ilustrativos revelam-se ao longo das 1000 páginas e cerca de 2000 entradas do *Glossary* e fazem dele um repositório único de vários saberes.

Sem nunca se falar especificamente sobre a operação linguística ‘empréstimo’ o *Glossary* tem como foco, muito simplesmente, a migração para o inglês, mais precisamente, Indo-Inglês, de segmentos linguísticos, sobretudo de origem indiana. Todo o processo de naturalização fonológica, morfológica e de adaptação semântica sofrido por essas transferências lexicais, induzidas por situações de contacto linguístico, se encontra registado acuradamente ao nível quer etimológico quer histórico. Todo esse labor filológico é clarificado e confirmado pelo recurso a um elenco imenso de obras de referência e pela inclusão de muitas e longas citações e de hipóteses etimológicas, assim como de narrativas a propósito, de modo que alguns verbetes, sendo verdadeiras monografias, ultrapassam de longe a microestrutura habitual do dicionário ou do glossário, mais parecendo a de uma enciclopédia. O referente e o nome que o identifica recebem a mesma atenção. Exemplo do que acabamos de dizer são as entradas, a duas colunas, dedicadas a: *Pardao* definido como “The popular

name among the Portuguese of a gold coin from the native mints of Western” com 5 páginas, *Typhoon* (A tornado or cyclone-wind) com 3,5 páginas, *Elephant*, um verdadeiro tratado etimológico de 3,5 páginas, ou a *Compound* que é definido como “The enclosed ground, whether garden or waste, which surrounds an Anglo-Indian house” com 3 páginas. A preocupação de Yule em juntar o nome à coisa que ele denomina leva-nos a reconhecer os autores do *Glossary* como seguidores *avant la lettre* da escola germânica ‘*Wörter und Sachen*’. Com todas as imperfeições que o linguista moderno possa apontar no método e na execução do *Glossary*, facto é que, até os nossos dias, ainda é considerado como uma obra lexicográfica séria para o estudo do Indo-Inglês, segundo o linguista indiano, Braj B. Kachru em 1980 “The only existing dictionary which closely approximates serious lexical work in Indian English” (Teltscher, 2013, p. xxi).

Resumindo, há particularmente três aspetos em relação a este glossário que desejamos destacar: o rigor e espírito científicos que os autores procuraram observar, e em grande parte conseguiram imprimir no texto, a extensa documentação compulsada e, de igual relevo, o interesse despertado pela sua leitura, que mais não é do que um dos propósitos (bem alcançado, a nosso ver) de Yule, como se regista no Prefácio (1903, p. x): “My first endeavour in preparing this work has been to make it accurate; my next to make it - even though a Glossary -interesting”.

2. AS FONTES DE ORIGEM PORTUGUESA DO *GLOSSARY*

Os autores do *Glossary* eram detentores de uma grande erudição representada no conhecimento dos clássicos, das obras dos sé-

culos XVII e XVIII, literárias ou científicas, de narrativas de viagens e de conquistas, por um grande conhecimento acerca do Oriente e, principalmente, por uma enorme curiosidade por tudo quanto com este se relaciona. Este interesse e informação foram, muito provavelmente, adquiridos aquando da sua permanência na Índia. A nível linguístico, como anteriormente se referiu, o saber era particularmente de Burnell, cujo conhecimento de sânscrito se somava ao de árabe, tibetano, copta e ao das línguas do sul da Índia. Todos esses aspetos se reconhecem nas cerca de 700 obras, abundantemente referidas ou citadas ao longo do *Glossary*. Em relação às de referência portuguesa registam-se 23 autores com 31 obras, e ainda outros títulos sobre temas relacionados com a presença portuguesa fora da Europa, sobretudo no Oriente, cujos autores ou não são portugueses ou escreveram em inglês. Consta-se ainda, na lista bibliográfica, que sobre algumas obras ou autores portugueses, como é o caso de *Archivo Português Oriental*, *Colloquios* ou *Lendas da Índia*, são tecidos comentários, sobre aspetos positivos, negativos e falhas, e acerca do valor do seu conteúdo.

Por que um recurso tão frequente e a tantos textos portugueses? A resposta é dada pelos autores; primeiro, ao apontarem para o papel fulcral do português como língua franca, principalmente no século XVI e depois, também nos séculos XVII e XVIII; segundo, porque estavam interessados não apenas em registar o discurso quotidiano, como era o objectivo inicial do *Glossary*, mas desejavam ir mais além recorrendo, para tal, a uma vasta gama de literatura relacionada com a Índia a todos os níveis do conhecimento. A valia do português como língua franca nas costas de África e, sobretudo na Ásia, durante a ocupação portuguesa e mesmo mais tarde, quando holandeses e ingleses os substituíram, não oferece quaisquer dúvidas para Yule. O dialecto Indo-Português afinal não fora apenas usado na comunica-

ção entre os europeus e os locais, mas igualmente entre os europeus de diferentes nacionalidades. A necessidade de saber português para poder negociar nessa parte do Oriente faz-se sentir até ao início do século XIX (Lopes, 1969, p. 31-117). Essa relevância do português como língua franca no Oriente durante esses três séculos será mais tarde magistralmente confirmada e documentada por Lopes em 1969. Segundo este, o século XVI que no Oriente era apenas português, irá ultrapassar o tempestuoso século XVII e chegar aos inícios do séc. XIX, acabando por sucumbir aos alvares da hegemonia política e comercial da Inglaterra.

Essa função da língua portuguesa como veículo de comunicação é ainda confirmada por 41 outras fontes estrangeiras onde se fala do uso do português no Oriente (Lopes, 1969, p. 31-117). Unicamente a título de exemplo e para se ter uma ideia da autoridade e credibilidade dadas por Yule às fontes portuguesas, fez-se uma contagem global de uma série de referências de uma fonte portuguesa nos verbetes, nomeadamente dos termos *Portuguese*, *Port. port.* e dos nomes dos autores portugueses mais conhecidos: R. Bluteau, Camões, D. Barbosa, Garcia de Orta e João de Barros. Isso mostrou que eles ocorrem em quase 60% dos verbetes. Esse mesmo tipo de contagem se poderia fazer para autores como: Diogo do Couto, Garcia de Resende, Fernão Lopes de Castanheda, Gaspar Correa, Fernão Mendes Pinto, ou para obras como *Roteiro de Vasco da Gama* etc. cuja frequência de citação ou de referência nos atesta como se revestiram de uma importância seminal na elaboração do *Hobson-Jobson*. Note-se, todavia, que Yule não se coíbe de assinalar também erros, negligências ou mesmo falhas, como mais adiante se faz notar no quadro que nós próprios criamos.

3. COLÓQUIOS DOS SIMPLES E DÉCADAS DA ÁSIA NO *HOBSON-JOBSON*

Quadro 1 – Edições compulsadas por Yule e por nós

Título	Autor	Por Yule	Por nós
<i>Colóquios dos simples e drogas he cousas medicinais da Índia e assi dalgũas frutas achadas nella onde se tratam algũas cousas tocantes a medicina, pratica, e outras cousas boas pera saber</i>	Garcia de Orta	2ª Edição, por F. Ad. de Varnhagen. Lisboa 1872, de acordo com a 1ª edição impressa em Goa em 1563, por Joannes de Endem	Edição dirigida e anotada por Francisco M. de Melo Breyner, Conde de Ficalho. Lisboa: Imp. Nac. 1891-1895, 2 vols. www.archive.org
<i>Da Asia de João de Barros e Diogo do Couto, dos feitos, que os Portuguezes fizeram no descobrimento, e conquista dos mares, e terras do Oriente</i>	João de Barros	Sobretudo a edição em 24 vols. com Diogo do Couto. Lisboa, na Regia Officina Typografica 1777-1778	A mesma que Yule consultou, mas em linha www.archive.org

Fonte: elaborado pela autora

Relativamente ao *Hobson-Jobson* servimo-nos da quarta edição, feita em Nova Deli em 1984; concretamente essa é uma reimpressão da edição de 1903 feita em Londres pelo orientalista William Crooke; usamos ainda a edição digital, disponibilizada pela *Digital Dictionaries of South Asia* <http://dsal.uchicago.edu/dictionaries/hobsonjobson/index.html>; essa última segue a edição de 1903 e foi objecto de uma última atualização em janeiro de 2006.

3.1. Análise do contributo de obras e autores portugueses para o *Hobson-Jobson*

Selecionaram-se os autores Garcia de Orta e João de Barros não só por serem os mais referidos por Yule, mas também por representarem duas obras de carácter, teor e estrutura bem diferentes. Esse último aspecto não implicou uma metodologia diferente relativamente à apreciação dos dados, que refletem o aproveitamento e a forma como Yule se serviu dos referidos autores e respectivas obras. Os passos, que a análise propriamente dita procurou seguir, foram os seguintes:

1. levantamento dos verbetes que mencionam o nome dos autores ou da obra seleccionados;
2. levantamento dos verbetes onde há citações das obras já referidas;
3. levantamento dos verbetes onde Yule se pronuncia criticamente, de modo positivo ou negativo, quanto ao merecimento do autor ou relevância da obra.

Partindo dos levantamentos indicados elaboramos o quadro 2:

Quadro 2 – Levantamentos

Autor	Verbetes	Citações	Observ. sobre o autor
Garcia (de Orta, Orta)	114	103	16
(João de) Barros	166	149	14

Fonte: elaborado pela autora

No quadro 2 e relativamente às menções e citações, inseridas nos verbetes apresentados e contados grosso modo, é evidente que as citações são em menor número em qualquer um dos autores. A leitura dos levantamentos nesse último quadro permite-nos fazer várias constatações e tecer algumas considerações. Assim, registramos que qualquer um dos autores pode ser mencionado sem que, obviamente, seja citado.

1. Relativamente a Garcia de Orta, verifica-se que o autor pode ser referido com a indicação do nome completo ou de partes do nome ou até mencionando, por extenso ou de forma abreviada, a obra em questão: ex.: Garcia, Garcia de Orta, De Orta, Colloquios ou Coll.

Exemplo do último caso, que é bastante mais raro, é o verbete *Pateca*. Acontece, porém, que no interior do verbete sempre aparece o nome Garcia e Garcia de Orta.

Apenas como curiosidade: no verbete *Coco-de-mer* a indicação do autor é dada em latim: “uti Garcias ab Orta persuadere voluit”.

2. Algo semelhante acontece com João de Barros, citado quase sempre por ‘Barros’ e raramente por ‘De Barros’ e, por exemplo, nos verbetes *Sabaio* por ‘João de Barros’ e *Umbrella* por ‘John de Barros’. É interessante verificar que Yule, no verbete *Sabaio*, se refere a J. de Barros como *our João de Barros*, o que denota um certo grau de conhecimento e de aproximação com esse autor, como adiante se verá.

A questão das citações é um pouco mais confusa. De um modo geral, sempre que se menciona o nome de um autor é incluída uma data e/ou uma citação como comprovativo. Por vezes, apenas se reproduz a forma de uma palavra como no caso de *Coromandel*, nome geográfico na costa sul da Índia mencionado por Garcia e por Barros nas formas *Choromandel* e *Charamandel*, respectivamente. Nenhum desses autores nem os muitos outros que incluem a palavra avançam uma hipótese explicativa plausível acerca da etimologia ou da gra-

fia do referido nome. Pensa-se que terá tido a sua origem em *Chola*, nome do povo que habitava na zona, porém, fica-se no plano das hipóteses. No entanto, parece que a forma mais antiga é, segundo Yule, a mencionada em 1497 no *Roteiro de Vasco da Gama*, publicado em Lisboa apenas em 1861, que apresenta a forma *Chomandarla*. Também na entrada dedicada ao fruto *Manga / mango* há várias referências ao português ou aos portugueses, sendo Garcia de Orta apenas uma vez citado, se bem que na mesma entrada as citações de outros autores ou obras sejam inúmeras.

Na entrada *Malabathrum*, Yule (4. ed., 1984, p. 543) diz-nos que, sem qualquer dúvida, esse termo clássico mais não é do que “uma folha seca ligada a várias espécies de canela”. E, embora teça um elogio a Garcia de Orta, faz-lhe igualmente um reparo. Na entrada *Umbrella* há uma menção a João de Barros, a qual, na realidade, pertence à entrada *Sombreiro*, Yule (4. ed., 1984, p. 851), razão pela qual não se encontra qualquer citação de Barros na referida entrada *Umbrella*. Outras vezes o autor é ‘mencionado’ pela negativa, isto é, não se menciona quando tal se espera; é o caso de *Custard-appel*, Yule (4. ed., 1984, p. 284) na respectiva entrada “The custard-apple is not mentioned by Garcia de Orta (1563)”. Na verdade, tendo-se em conta que o fruto *Custard-appel*, nome inglês da anona, parece ter sido levado para a Índia por portugueses e daí o nome dado no Malabar, “the *Parangi i.e. Firingi* or Portuguese”, seria de esperar uma referência em Garcia de Orta. No capítulo *Anones*, Ferrão (2020, p. 59-64) confirma o papel dos portugueses na introdução da anona e suas espécies em África e na Índia. Caso semelhante acontece no verbete *Myrobalan* [mirabólanos] (Yule, 4. ed., 1984, p. 607- 610) o nome de uma glândula e designação genérica de vários frutos secos, procedentes da Índia com vários aproveitamentos especialmente na farmacopeia. Nesse verbete menciona-se seis vezes Garcia de Orta, aliás como seria de

esperar visto tratar de matéria usada em preparações farmacêuticas, no entanto, sem que haja, uma citação do *Colóquio* 37 (150-167), que versa detalhadamente sobre os *mirabólanos*, suas espécies e seus usos. Também no verbete *Abada* ‘rinoceronte’ (Yule, 4. ed., 1984, p. 1) J. de Barros é mencionado mais do que uma vez sem, no entanto, ser citado. Obviamente, o facto de ser referido não implica o registo de um extracto das obras em causa.

Quanto ao título das Obras referidas ou mencionadas, Yule indica com mais frequência o nome do autor do que o da obra. Relativamente à forma abreviada Coll. esta pode referenciar obras de outros autores e não apenas Garcia de Orta, daí a necessidade sentida por Yule de indicar igualmente o nome do autor. A menção por extenso, *Colloquios*, apresenta apenas sete ocorrências em todo o *Glossary*, mas o nome do autor é igualmente mencionado. Com toda esta especificação não há hipótese de equívoco. Yule reserva o nome *Década* para a obra de Bocarro. Ainda relativamente às *Décadas da Ásia*, se a referência se faz às primeiras, o autor indicado é, naturalmente, J. de Barros e as datas apresentadas são 1552, 1553 e 1563; usando-se (Diogo do) Couto a partir da quarta ou para o *Soldado Prático*. Isso embora na bibliografia Yule diga que se serviu especialmente da edição de 1777-1778.

Em relação à forma como se cita o excerto, quer de Garcia de Orta quer de Barros, verificam-se os seguintes modos:

1. O excerto é invariavelmente traduzido para inglês, embora possa acontecer que o item em causa seja igualmente assinalado no original português.

Exemplos em relação a Garcia de Orta:

Verbetes: *caste*, *mustter*, *parsee* // port.: **casta** [Colóq. 54], **mostra** [Colóq.4], **mercadores de boticas** [Colóq. 54]

Exemplos relativos a João de Barros:

Verbetes: *alligator*; *assegay*; *mustees*, *mestiz* // port. **lagartos** [I, iii, cap.8], **azagaias** [I, iii, cap.1], **mistiços** [I, ii, cap.1].

2. Muito raramente o extracto é oferecido no original português. Em todo o *Glossary* apenas se encontraram três citações retiradas dos *Colóquios* e mantidas no original nos verbetes a saber: *Acheen* [Col.Ode], *Palmyra* [Col.14] e *Zedoary* [Col.57]; no primeiro caso trata-se de uma estrofe da Ode dedicada por Camões a G. de Orta, publicada no primeiro volume e considerada o primeiro trabalho publicado de Camões. Interessante é o facto de que quando se citam *Os Lusíadas* se regista, na maior parte das vezes, o texto original seguido pela sua tradução feita ou por Burton (1880) ou por Aubertin (1878). O mesmo acontece umas duas vezes com o *Roteiro de V. da Gama*. J. de Barros apenas é traduzido em quatro verbetes a saber: *Caffer*, *caffre* [Dec.II,i, cap.1], *Coir* [Dec.III,iii, cap.7], *Tiger* [Dec.II,vi, cap.1] e *Veranda* [Dec.II, vi, cap.3].

A partir do quadro 2 com os levantamentos referentes aos dois autores em apreço e de acordo com as considerações feitas no *Glossary*, elaboramos um terceiro e quarto quadros. O quadro 3 compreende as menções honrosas e elogios sobre o autor ou a obra; o quadro 4 mostra as ausências e reparos detectados. As primeiras partes dos quadros são dedicadas a Garcia de Orta e as segundas a João de Barros.

É evidente que muitas vezes, sem se citar, se remete simplesmente, para um dos autores, todavia, para além das menções acima assinaladas (*cf.* frase em epígrafe e acerca das obras de referência portuguesas no ponto 1.1), salientam-se ainda vários louvores quer a G. de Orta quer a J. de Barros.

Quadro 3 – Menções honrosas ou elogiosas diretas e subentendidas

Autor / obra	Segmento	Entrada do Glossary
<i>Colloquios</i>	A most valuable book , full of curious matter and good sense	<i>Obras citadas</i> p. xxxiv/v
G. de Orta	G.de O. who first described the virtues of this fruit in the treatment of dysentery	Bael (marmelos)
Garcia de Orta	and expressly alleged, by the judicious Garcia de Orta	Bengal
De O.	De O., whose judgments are always worthy of respect	Catechu
De Orta	Even the sagacious De Orta	Comar
Garcia	is a mistake as indeed the sagacious Garcia also pointed out	Mace
<i>Colloquios</i>	The matter is treated, with his usual lucidity and abundance of local knowledge ,	Malabathrum
Garcia de Orta	as the estimable Garcia de Orta says	Maldives
G.de Orta	G.de Orta, who is trustworthy , says explicitly that	Myrobalan
Garcia de Orta	This root was recognised by the famous Garcia de Orta	Putchock
G. de Orta	G. de Orta goes at length into this subject	Tabasheer
Barros	The more common etymology is that which is given by Barros, (<Spanish word coco ..)	Coco
De Barros	the usually well-informed De Barros	Deccan
De Barros	Since De Barros, Malabar occurs almost universally .	Malabar
Barros	Pegu is the usual Portuguese form, as in Barros, and so passed to us	Pegu
De Barros	the intelligent statement of De Barros	Sabaio
Barros	Lest I am doing the great historian wrong	Tiger

Fonte: elaborado pela autora

Quadro 4 – Reparos negativos

Autor	Segmento	Entrada do <i>Glossary</i>
Garcia de Orta	The proper form of this name appears to be <i>Male-dīva</i> ; not , as the estimable Garcia de Orta says, <i>Nale-dīva</i>	Maldives
Garcia	as he repeats the very errors which Garcia dissipates	Malabathrum
Garcia de Orta	Yet the whole of what is said by Garcia de Orta is inconsistent with this	Pateca
De O.	De Orta speaks of [...] but his informants misled him.	Putchock
Garcia de Orta	Garcia de Orta does not seem to have been aware of this history and he derives the name from <i>Sāhib</i>	Sabaio
Barros	This and the next passage compared show that Barros was not aware that <i>Cospetir</i> and <i>Gajpati</i> were the same	Cospetir
Barros	but of course the two first [explanations] are quite wrong . <i>Iniza Malmulco</i> , as Barros here writes it, is Ar. <i>An-Niẓām ul Mulk</i> , “The Administrator of the State,” not from P. <i>neza</i> , “a spear.”	Cotamulco
De Barros	It is difficult to discover what has led astray here the usually well-informed De Barros	Deccan
De Barros	Barros [...] but takes no notice of the great river, so far as we are aware, in any part of his history	Godavery
Barros	The fact that the poet copies the misprint or mistake of Barros in <i>Asaboro</i>	Mussendom
João de Barros	As our João de Barros could not get true information of these things, he confounded the name of the Hindu <i>Savay</i> with that of Çufo (? Yūsuf) Adil Shāh	Sabaio
De Barros	The explanation given by De Barros and Couto is altogether fanciful	Singalese
De Barros	In this passage De Barros goes unusually astray	Zirbad

Fonte: elaborado pela autora

Tomadas essas observações conjuntamente, constata-se que Barros é mais citado do que Orta; no entanto, ressalta o facto de Yule ter sido mais pródigo em juízos de valor positivos em favor daquele último. Relativamente a falhas ou incorreções o levantamento aponta para um maior número em Barros; tal não deve causar estranheza tendo presente o maior número de ocorrências referentes a esse autor. Diga-se, porém, que os aspectos negativos atribuídos a Orta e a Barros se podem considerar como “falhas menores” comparados não só com o largo aproveitamento feito desses autores e dos outros, não tratados neste texto, como com o volume das obras em análise e das outras não consideradas. Atrás, no quadro referente a falhas em João de Barros, Yule emprega “our João de Barros” o que denuncia e acentua o grau elevado de conhecimento e familiarização com a obra deste, manifestados por Yule.

Creemos ainda que as observações, apoiadas em textos e testemunhos de outros autores da época, e tecidas por Yule e Burnel nas *Introductory Remarks*, acerca do papel do Indo-Português, como língua franca, nos séculos XVI e XVII e talvez ainda no século XVIII, assim como o uso feito e comentado dos dois autores seleccionados e trabalhados neste artigo e de alguns mais como Gaspar Correia, Fernão Lopes de Castanheda, para citar apenas os mais mencionados, são matéria suficiente para justificar o epíteto *alambique* dado por Yule ao português na gênese do Indo-Inglês (1903, XVIII-XIX). A palavra *alambique* provém do árabe hispânico, que, por sua vez, teve a sua origem no grego. Alambique é um equipamento usado, sobretudo na destilação de certos líquidos, feita por aquecimento ou fermentação, resultando um líquido semelhante, mas nunca igual, uma vez que a passagem pelo alambique deixou as suas marcas. O mesmo terá acontecido na formação do Indo-Inglês atendendo à função de filtro do Indo-Português, entre nativos e não nativos mesmo após a derrota

dos portugueses e substituição primeiro pelos holandeses e depois pelos ingleses. Tal como o líquido passado por um alambique adquire outros traços também as línguas Indo adquiriram marcas deixadas pelo Indo-Português, enquanto língua franca naquela zona da Índia.

SÍNTESE CONCLUSIVA

Tendo em conta a quantidade e gama de informação recolhida nos textos, sejam eles portugueses ou não, redigidos nas mais diversas línguas europeias e orientais, trabalhada apenas por dois lexicólogos-lexicógrafos, e, obviamente, sem a bateria de auxílio informático, como hoje dispomos, a tarefa levada a cabo por Yule e Burnell afigura-se nos gigantesca. Apesar de ser muito abrangente e de apresentar entradas extremamente completas com anotações de todo o tipo, nota-se a ausência de certos conceitos de tipo teológico e / ou filosófico, mesmo tendo presente o facto de a obra ter sido levada a cabo em Inglaterra, durante a época vitoriana (1837-1901).

Procuramos pôr em evidência o contributo dado pelas duas obras portuguesas em apreço na construção do *Glossary Hobson-Jobson* e, assim, verificou-se que aquelas e as muitas outras que não foram objecto da nossa pesquisa direta, se revestiram de seminal importância na economia desse glossário de Indo-Inglês. A diversidade de assuntos focados nas obras aqui seleccionadas ao ser lexical e etnograficamente utilizada pelos autores, mostra, de forma ajustada e inequívoca, a importância dos factores exógenos no percurso e grau de integração do léxico importado e transformaram o *Glossary* num produto único. Muitas das explicações etimológicas oferecidas no glossário estão escoradas em textos portugueses, e não só, de vária natureza e de várias épocas, além disso, são confirmadas através não de uma citação, mas de um manancial de citações dos mesmos. Só um

conhecimento profundo dessas obras e de muitas outras não portuguesas permitiu aos autores a construção de um glossário como o presente. Como já afirmamos (Augusto, 2006) a propósito do recurso a Rafael Bluteau no *Glossary*, as citações, invariavelmente vertidas para inglês, representam os segmentos que os autores do glossário consideraram fulcrais por veicularem valor linguístico, sócio-histórico, botânico, geográfico, etnográfico ou outro. Deste modo, o contributo linguístico, histórico e textual das obras portuguesas revelou-se não só fundamental como imprescindível, ou melhor, obrigatório, no processo de argumentação presente no *Hobson-Jobson*; daí talvez também a elaboração da **Nota B**, sobre o Português com observações se bem que elementares, que se destaca por ser a única alusão gramatical ao emprego de uma língua. Isso e o facto de nos séculos XVI e XVII e muito provavelmente ainda no século XVIII, o Indo-Português, como acima afirmamos, ter operado no sul da Índia, como língua de contacto na esfera social e ou comercial, fez dele um elemento incontornável no panorama linguístico da Índia com uma grande incidência na emergência e na formação do Indo-Inglês.

Em 2014, comparamos o *Hobson-Jobson* com dois outros dicionários do mesmo tipo produzidos na época: o inglês *Dictionary Phrase and Fable*, do Reverend E. Cobham Brewer (1870) e o holandês *Woor-denschat: verklaring van woorden en uitdrukkingen*, de Taco Hajo De Beer e Eliza Laurillard (1893-1899). Chegamos então à conclusão que o *Hobson-Jobson*, apesar das suas falhas ultrapassar de longe os outros dois, quer em quantidade de entradas quer no tratamento minucioso e exaustivo das mesmas.

Resta dizer que o merecimento das abundantes citações, sejam de autores portugueses, sejam de outros, de cariz linguístico ou não, longe de tornarem o *Glossary* fastidioso transformaram-no num monumento rico ao nível lexicográfico, histórico e etnográfico e fazem

com que a sua leitura seja não apenas informativa, mas igualmente fascinante; como diz Dalby (1998), o *Hobson-Jobson*, por produzir a história das palavras importa para a lexicografia inglesa, mas também para outras línguas como o português ou o francês.

NB: Agradeço a José Bettencourt e a Jef van Egmond a leitura atenta do manuscrito; porém, os possíveis erros que subsistam são da minha responsabilidade.

REFERÊNCIAS

AUGUSTO, M. C. Dom Rafael Bluteau e Henry Yule – a função do Vocabulário Portuguez e Latino no Hobson-Jobson. In: THIELEMANN, W. (ed.) **Congresso Século das Luzes – Portugal e Espanha, o Brasil e a Região do Rio da Prata**. Frankfurt am Main: Teo Ferrer de Mesquita, 2006. p. 95-116.

BARROS, J. de, COUTO, D. do. **Décadas da Ásia**: Da Ásia de João de Barros e de Diogo do: dos feitos que os portugueses fizeram no descobrimento dos mares e terras do Oriente. Volumes I (1552), II (1553), III (1563) e IV (1615). Nova ed. Lisboa: Na Regia officina typografica, 1552-1615/ 1777-1788. Disponível em: <http://www.archive.org/search.php?query=jo%C3%A3o%20de%20barros%20AND%20mediatype%3Atexts>. Acesso em: 10 out. 2021.

BLUTEAU, R. **Vocabulário portuguez e latino**. 4. v. Coimbra: Real Collegio das Artes da Companhia, 1712 e 1728. Disponível em: <https://purl.pt/13969/4/>. Acesso em: 10 out. 2021.

DALBY, A. **A guide to world Language Dictionaries**. London: Library Association Publishing, 1998.

DIEZ, F. **Etymologisches Wörterbuch der romanischen Sprachen**. Paderborn: Salzwasser-Verlag GmbH.1869

FERRÃO, J. E. M. **Le Voyage des Plantes & Les Grandes Découvertes**. Paris: Éditions Chandeigne, 2020.

LITTRÉ, É. **Dictionnaire de la langue française**, (1873 à 1877). Dernière actualisation: Avril 2018. Disponível em: <https://www.littre.org/>. Acesso em: 10 out. 2021.

LOPES, D. **Expansão da Língua Portuguesa no Oriente nos séculos XVI, XVII e XVIII**. 2. ed. Rev. e aumentada por Luís de Matos. Porto: Portucalense Ed., 1969.

ORTA, G. de. **Colóquios dos Simples e Drogas da Índia**. [Goa, 1563], edição preparada e anotada por: Conde de Ficalho, 2 vols. Lisboa: Academia Real das Ciências de Lisboa / Imprensa Nacional, 1891-1895. Disponível em: <http://www.archive.org/details/coloquiosdossimp01ortauoft>. Acesso em: 10 out. 2021.

TELTSCHER, K. The Floating Lexicon: *Hobson-Jobson* and the *OED*. In: SENGUPTA, I.; ALI, D. (Eds.). **Knowledge Production, Pedagogy, and Institutions in Colonial India**. London: Palgrave Macmillan, 2011. p. 43-57.

TELTSCHER, K. **Hobson-Jobson: The Definitive Glossary of British India**. Oxford: Oxford University Press. 2013.

YULE, H.; BURNELL, A. C. **Hobson-Jobson: A Glossary of Colloquial Anglo-Indian Words and Phrases, and of Kindred Terms, Etymological, Historical, Geographical and Discursive**. 4. ed. New Delhi: Munshiram Manoharlal Publishers, 1984. Disponível em: <https://openlibrary.org/books/OL6903772M/Hobson-Jobson>. Acesso em: 10 out. 2021.

DICIONÁRIO MONOLÍNGUE DE LÍNGUA ESTRANGEIRA: A FACE VISÍVEL E A FACE INVISÍVEL DAS PALAVRAS

Álvaro David Hwang
Maria Cristina Parreira

INTRODUÇÃO



O presente trabalho resulta do interesse por questões metalexigráficas relativas a dicionários monolíngues de língua francesa, não somente enquanto pesquisadores em Ciências do Léxico, mas também como usuários que recorrem, de forma assídua, a essa importante ferramenta de aprendizagem, tanto em sala de aula quanto em diversas situações de estudo ou pesquisa, buscando, na obra lexicográfica, informações que, por um lado, levem a encontrar respostas para o que motivou uma consulta e que, por outro, promovam, no usuário, o desenvolvimento de uma aprendizagem autônoma, reafirmando, assim, o papel de primeiro plano que os dicionários têm no ensino-aprendizagem de línguas.

Este estudo procura tratar de questões de interesse do usuário que abre um dicionário para aprender uma determinada língua, procurando encontrar, no dicionário monolíngue de língua estrangeira, em particular, uma ferramenta transformadora de sua experiência com o idioma, com base, sobretudo, nos aspectos da descrição que

revelam o potencial da obra lexicográfica para o desenvolvimento de sua capacidade de produção na língua.

Quanto a esse aspecto, é importante lembrar que os exemplos têm um papel essencial para o desenvolvimento da capacidade de codificação do usuário e que essa é uma razão importante para que, conforme nos lembra Calderón Campos (1994), pareça ser consensual o entendimento de que eles são peças centrais do enunciado lexicográfico. De fato, eles são importantes para a apreensão das propriedades léxico-semânticas da unidade lexical descrita, o que permitirá tanto a compreensão quanto o emprego adequado de uma palavra. Eles também são importantes para mostrar seu comportamento sintático-semântico, pois a categoria gramatical de uma unidade lexical não fornece nenhuma informação precisa sobre isso, conforme afirmam Fradin e Marandin (1979), que nos lembram ainda que, se um dicionário não apresentar exemplos que forneçam essas informações sintático-semânticas, elas não serão encontradas em lugar algum. E isso tem um forte impacto nas condições dadas pela obra lexicográfica para o desenvolvimento da capacidade de codificação na língua por parte daquele que consulta o dicionário, dado que, muitas vezes, o sentido está intimamente ligado às condições sintáticas de uso de uma unidade lexical, o que torna, portanto, as condições sintático-semânticas de emprego dessa unidade uma informação tão importante quanto o próprio sentido. Essa informação permitirá ao usuário tanto observar as coerções sintagmáticas que se impõem à produção de enunciados na língua descrita quanto avaliar as (im)possibilidades sintagmáticas que decorrem dessas coerções.

Dentre outras funções importantes, os exemplos são um lugar privilegiado para informar também sobre os coocorrentes mais frequentes da palavra descrita, dando ao usuário a oportunidade de aperfeiçoar sua capacidade de selecionar e usar apropriadamente uni-

dades lexicais em relação de compatibilidade combinatória (*cometer um crime*/**fazer um crime*). Conforme afirma Rey (1989, p. 64), os exemplos têm, na verdade, uma realidade complexa e as informações que eles podem vincular são de natureza variada. Por serem considerados, tanto pela teoria como pela prática lexicográfica dos últimos anos, como elementos centrais do enunciado lexicográfico, Mederos (1994) afirma, com razão, que sua presença nos dicionários gerais é cada vez mais necessária, sendo sua ausência nos dicionários pedagógicos imperdoável.

De forma mais ampla, a riqueza dos exemplos, em termos da potencialidade de informações que eles veiculam, foi muito bem traduzida por Rey-Debove (2005, p. 18) quando, ao falar do estatuto e função do exemplo no dicionário, ela diz que “o uso da palavra em discurso pode sugerir *predicados implícitos complementares*: a passagem da língua ao discurso é fundamental para se conhecer a palavra” (grifo nosso). Sabemos, conforme já afirmava a mesma autora (Rey-Debove, 1971), numa época em que se consolidavam os primeiros trabalhos metalexográficos franceses, que o enunciado lexicográfico, do ponto de vista da articulação dos seus elementos constitutivos enquanto relação predicativa, se constitui de um conjunto de predicados (*gato se pronuncia de tal forma, gato é um substantivo, gato significa tal coisa etc.*). Esse conjunto de predicados fornecidos sobre a entrada dá ao enunciado lexicográfico a aparência formal de uma sequência fechada de informações, formando blocos textuais mais ou menos compactos que refletem, conforme nos lembra a mesma autora (Rey-Debove, 1971), a estrutura real da obra lexicográfica que, por sua natureza, não permite uma leitura contínua. Nesse sentido, sabemos que as remissivas explícitas que permeiam o verbete são uma forma de *corrigir* uma visão estanque do enunciado lexicográfico e que outros vínculos podem ser estabelecidos indiretamente entre diferentes verbetes,

como acontece com as remissivas implícitas, por exemplo, que são percursos que o próprio usuário pode estabelecer conforme se depara, no interior do próprio texto lexicográfico, com um item lexical que o leva, por desconhecimento ou dúvida, a uma nova consulta.

Se nos detivermos, no entanto, com mais atenção à afirmação acima citada, podemos concluir que os exemplos são um espaço privilegiado que, no enunciado lexicográfico, oferece ao usuário a possibilidade de construção de informações novas sobre a língua, ou seja, de formular *predicados implícitos* que, muitas vezes não predizíveis, podem ser construídos pelo usuário, conforme sua capacidade de observação, reflexão e diálogo com o enunciado lexicográfico, o que o distingue de tantos outros usuários que podem também construir tantos outros predicados implícitos não predizíveis. Embora interessante, essa questão específica não será tratada aqui, mas ela vem reforçar o entendimento do lugar de primeiro plano que os exemplos têm no enunciado lexicográfico e terão neste presente trabalho, ainda que, na análise a ser feita, os enunciados definicionais sejam abordados de forma igualmente importante, o que vem a reafirmar o papel desses dois elementos como alicerces do enunciado lexicográfico. O presente trabalho busca observar como se dão as relações entre esses paradigmas de informação, na perspectiva do que estamos chamando de *face visível* e *face invisível* das palavras, conforme será explicado em seguida.

O tratamento dos aspectos relativos ao que estamos chamando de *face visível* e de *face invisível* das palavras tem relação com uma questão bem mais ampla e crucial na gênese ou na produção de uma obra lexicográfica: o controle das informações fornecidas em diferentes enunciados lexicográficos e a rede implícita de relações que elas estabelecem entre si, promovendo ou comprometendo a coesão interna do discurso lexicográfico como um todo. Para o lexicógrafo, é uma

tarefa de grande complexidade e um desafio permanente controlar as informações fornecidas em toda sua rede tentacular de relações visíveis e, em muitos casos, invisíveis que permeiam as tramas do discurso lexicográfico como um todo, uma vez que elas se expandem para além do espaço topológico do enunciado lexicográfico. Isso decorre da evidência de que a completude do verbete (*face visível* da palavra) não se realiza somente nele, mas dialoga e interage com outros enunciados lexicográficos, sem que, muitas vezes, essas relações implícitas no interior do discurso lexicográfico tenham sequer sido programadas para serem visíveis ao usuário (*face invisível* da palavra).

É muito importante dizer que o presente trabalho não tem natureza crítica, no sentido de que a descrição dos fatos observados não tem a ambição de apontar insuficiências na descrição lexicográfica, mas sim reforçar a ideia de como uma *face invisível* das palavras pode ser observada na obra lexicográfica. Isso nos permite dizer, por uma metáfora talvez feliz, que a riqueza do discurso lexicográfico está na possibilidade constante de erupção de informações em diferentes episódios de consulta e leitura, sejam elas informações explicitamente dadas de forma dispersa no discurso lexicográfico ou informações que, conforme dissemos anteriormente, podem ser implicitamente construídas pelo usuário.

Tendo como interesse principal as informações sintático-semânticas que permitem a produção de enunciados na língua, em conformidade com as coerções sintagmáticas que se impõem ao uso de determinadas unidades lexicais, a análise nos permitirá observar primeiramente que há, no interior do discurso lexicográfico, uma circularidade de informações que promove, entre diferentes enunciados lexicográficos, uma coesão interna e certa permeabilidade. Essa circularidade deve ser entendida no sentido de que existem, entre verbetes, zonas de contato que revelam conexões implícitas de informação que

lhes dão certo grau de coesão, à medida que as mesmas informações transitam entre eles. Nesse sentido, ela corresponde a um aspecto positivo do discurso lexicográfico, pois é um indicador de que há trânsito de informações entre enunciados lexicográficos, ou seja, uma cumplicidade entre eles que cria efeitos de redundância pedagógica. Em Hwang (2023), são apresentados alguns aspectos desses efeitos graças à sistematicidade na forma de apresentação das informações.

O interesse principal desta pesquisa é observar como, em diferentes enunciados definicionais mais ou menos dispersos no discurso lexicográfico, encontram-se informações sintático-semânticas sobre a relação de coocorrentes privilegiados que não são fornecidas na própria descrição da unidade lexical nominal que figura espontaneamente nesses enunciados definicionais, revelando o que estamos chamando de *face invisível* das palavras. Nesse caso, esses enunciados definicionais dispersos na obra lexicográfica revelam uma situação peculiar em que a definição, dentro da perspectiva que aqui opõe a *face visível* e a *face invisível* das palavras, pode ser observada como exemplo *incógnito* que informa de maneira *acidental* sobre condições sintático-semânticas de funcionamento de uma unidade lexical: *incógnito*, porque ocorre fora do espaço topológico do enunciado lexicográfico destinado à descrição dessa unidade, o que não viabiliza o acesso a essa informação pelo usuário que, interessado em informar-se sobre as condições de uso da palavra, dirige-se apenas ao verbete que informa sobre ela; *acidental*, pois não se trata de um enunciado destinado a ilustrar as condições de funcionamento da unidade lexical constitutiva do enunciado definicional.

Esta pesquisa alinha-se a uma proposta que Pruvost (2005) e Martinez e Pruvost (2010) denominam de tripla investigação dicionarística, na qual é possível observar o uso do termo dicionarística em oposição ao termo lexicografia, segundo uma distinção conceitual

fundamental feita por Quemada (1987) entre esses dois termos. Para Pruvost (2005), essa tripla investigação tem grandes contribuições a trazer para a Lexicografia e, em particular, para a produção de obras lexicográficas de qualidade cada vez melhor. Apresentamos em seguida as linhas principais dessa proposta.

1. A TRIPLA INVESTIGAÇÃO DICIONARÍSTICA

Pruvost (2005) e Martinez e Pruvost (2010) propõem três procedimentos de busca e análise de informações, no interior do discurso lexicográfico, a respeito de uma determinada palavra. Esses procedimentos, que os referidos autores denominam de tripla investigação dicionarística, têm como objetivo principal a prospecção de todas as informações lexicográficas sobre uma unidade lexical, estendendo o aspecto arqueológico desse trabalho para além do perímetro do enunciado lexicográfico destinado à descrição dessa unidade. Segundo Pruvost (2005), os resultados dessa tripla investigação surpreenderiam tanto o usuário, pela riqueza de dados obtidos a respeito de uma palavra ou conceito, quanto o próprio lexicógrafo, pela quantidade de informações que, em alguns casos, nem ele teria suspeitado trazer à tona no discurso lexicográfico.

Outra vantagem importante dessa tripla investigação seria a oportunidade de, a partir das informações obtidas, avaliar melhor a qualidade técnica da descrição oferecida no enunciado lexicográfico destinado ao tratamento da unidade lexical *investigada* e, eventualmente, melhorar aspectos de sua descrição. A primeira investigação sugere a busca de uma mesma informação em obras lexicográficas diferentes, desde que sejam produtos, no entanto, de projetos lexicográficos análogos. Em seu artigo, Pruvost (2005) analisa o tratamento lexicográfico dado à palavra *norma*, nos dicionários *Le Petit Robert* e

Petit Larousse, o que lhe permitiu levantar aspectos que, desde a estruturação formal do verbete à organização dos sentidos e à exemplificação, mostram diferentes interpretações explícitas da palavra, assim como traços comuns entre elas.

A segunda investigação dicionarística consiste em buscar, numa mesma obra lexicográfica, todas as ocorrências de uma mesma palavra, o que pode ser feito sistematicamente quando se dispõe de uma versão digital do dicionário, pois isso permite sair ao encaixe de uma palavra em toda a obra, ou seja, no interior dos mais diferentes verbetes, entre os quais ela estabelece implicitamente uma rede variada e complexa de relações, por figurar tanto em enunciados definicionais e glosas explicativas quanto nos mais diferentes tipos de exemplos fornecidos.

Conforme diz o autor, uma rede de informações, que delinea uma face não diretamente visível das palavras, revela-se presente no corpo do dicionário. Ele ainda afirma que essa segunda investigação fornece uma espécie de tomografia dicionarística da palavra, dando dela imagens que permitem medir o que ela representa quando ocorre de forma *espontânea*, ou seja, fora das condições de *alta vigilância* a que ela está submetida quando é objeto de tratamento lexicográfico no próprio verbete que a descreve. A presente pesquisa inspira-se essencialmente nesse aspecto da tripla investigação dicionarística.

A terceira investigação procura obter uma radiografia concordancial da palavra, procurando rastrear, nas suas ocorrências em diferentes verbetes do dicionário, elementos de informação sobre a natureza semântica e sintática da palavra, que, conforme afirma o autor, podem, até mesmo sem intenção explícita dos dicionaristas, completar informações sobre traços sintático-semânticos da palavra por meio de relações de natureza concordancial, ou seja, que definem

contextos em que figuram tanto as palavras que a precedem quanto aquelas que a seguem. Além dessa preocupação no eixo sintagmático, a terceira investigação também implica o rastreamento de relações paradigmáticas, o que permite um monitoramento das coerências internas do discurso lexicográfico, no tecer das relações analógicas, sinonímicas e antonímicas entre diferentes enunciados lexicográficos. Em suma, a tripla investigação dicionarística representa uma forma de conciliação da *face visível* e da *face invisível* das palavras, tendo, portanto, muitas contribuições a trazer tanto para o desvendamento das riquezas inimagináveis do discurso lexicográfico quanto para o aperfeiçoamento do trabalho lexicográfico e, por conseguinte, para a melhoria da qualidade técnica das obras que resultam desse trabalho.

2. A FACE VISÍVEL E A FACE INVISÍVEL DAS PALAVRAS

A análise tem como objeto o dicionário *Le Petit Robert - Dictionnaire alphabétique et analogique de la langue française (Le Petit Robert, CD-ROM, 2011)*. Essa escolha se deveu ao fato de se tratar de uma obra de referência da língua francesa e ao fato de que esse dicionário está disponível em uma versão informatizada, em CD-ROM, o que facilita a busca, de forma exaustiva e completa, das ocorrências de uma determinada unidade lexical em diferentes enunciados. Isso permite, conforme foi dito, sair ao encalço de uma palavra e localizar todos os verbetes em que ela aparece. Apesar da possibilidade do tratamento exaustivo dessas ocorrências, a análise se restringe a fatos que permitem apenas ilustrar o que estamos entendendo como *face visível* e *face invisível* das palavras, conforme já foi comentado.

A seleção das unidades lexicais para fins de análise teve como ponto de partida a confrontação de informações entre enunciados lexicográficos que descrevem palavras pertencentes à mesma família morfossemântica. Primeiramente porque, no interior desse conjunto, já existe entre essas unidades uma forte coesão, por apresentarem entre si uma relação de identidade observável tanto no nível morfológico quanto no nível semântico. Na descrição lexicográfica, essa coesão, que se traduz por uma relação de parentesco entre elas, tem relação com o fato de que as definições morfossemânticas sejam muito frequentes na descrição desses subconjuntos lexicais, pois a natureza das relações léxico-semânticas que estruturam esses subconjuntos torna esse procedimento de definição a forma mais simples e mais econômica de definir (Rey-Debove, 1971). Nesse tipo específico de definição (ex.: *influent: qui a de l'influence*), o enunciado definicional (*qui a de l'influence*) tenta tratar da complexidade do definido pela retomada de sua base lexical (*influence*). Enquanto elemento de contato entre diferentes enunciados lexicográficos, essa retomada dá ocasião a ocorrências dessa base com coocorrentes privilegiados (ex.: *avoir + influence*), formando combinatórias lexicais que são de interesse para a presente pesquisa à medida que nelas intervêm coerções sintático-semânticas não predizíveis pelo usuário estrangeiro (*avoir de l'influence/*avoir influence*). A análise permitirá também observar essas questões fora da esfera de famílias morfossemânticas, pois as redes de relações implícitas entre enunciados lexicográficos vão além da esfera das relações de comunidade de sentido e forma.

Conforme foi comentado anteriormente, o interesse principal desta pesquisa é verificar se a descrição de uma unidade lexical nominal traz, nos exemplos fornecidos, informações sobre os mecanismos sintático-semânticos que regem as relações dessa unidade e os coocorrentes privilegiados com os quais ela figura em enunciados defini-

cionais dispersos no discurso lexicográfico, criando entre diferentes verbetes relações implícitas que, nos casos em que isso acontece, nos permitirão comentar aspectos da *face visível* das palavras. Nos casos em que isso não acontece, estão dadas as condições para falarmos em *face invisível* das palavras.

Tendo, conforme foi dito, interesse por aspectos da descrição que promovam a capacidade de produção na língua descrita e, em particular, pelas informações sobre as coerções sintático-semânticas que regem os funcionamentos de uma unidade lexical, o presente trabalho permitirá indiretamente avaliar a vocação do enunciado lexicográfico para a instrumentalização do usuário para o domínio dessas propriedades da língua. A análise terá como foco principal as coerções sintagmáticas que se impõem a combinações lexicais do tipo Verbo + Sub. como, por exemplo, *avoir + talent*, *avoir + courage*, *porter + préjudice* etc, pois nelas entram em jogo mecanismos de determinação ou de ausência de determinantes que, na verdade, representam dificuldade para quem aprende francês.

Na análise, as informações foram transcritas parcialmente, focalizando os elementos de informação pertinentes para a ilustração do que é tratado, com os cuidados necessários para que os recortes realizados na transcrição não trouxessem prejuízo para a legitimidade da análise. Nos casos de acúmulo paradigmático (ex.: *influent: qui a de l'influence, du prestige, du crédit*), apenas o paradigma em foco é transcrito (ex.: *influent: qui a de l'influence* ou *influent: qui a du prestige* ou *influent: qui a du crédit*). Os parênteses pontilhados [...] indicam a omissão de comentários semânticos, a fim de introduzir a transcrição do exemplo dado na descrição da unidade tratada.

2.1. A *face visível* das palavras

É possível notar que existe entre alguns verbetes um circuito fechado de informações sintático-semânticas que estabelece entre eles uma zona de contato, à medida que uma mesma informação circula entre eles. Em *talentueux, eusse: qui a du talent*, o mecanismo sintático-semântico da fórmula definicional é fornecido como exemplo na descrição de *talent: II. 2. [...] avoir du talent*, o que dá ao usuário a oportunidade de se informar, na própria descrição dessa unidade lexical, sobre coerções sintático-semânticas importantes para a produção de enunciados dessa natureza e para a avaliação das (im)possibilidades sintagmáticas que decorrem do conhecimento dessas informações (ex.: *avoir du talent/*avoir talent; qui a du talent/*qui a talent; ne pas avoir de talent/*ne pas avoir talent; avoir un talent exceptionnel/*avoir du talent exceptionnel etc.*). Além disso, a explicitação dessa informação (*face visível* da palavra) também lhe dá condições melhores para reconhecer tais mecanismos tanto em suas outras experiências de leitura quanto em enunciados definicionais que eventualmente se encontrem dispersos no discurso lexicográfico e nos quais essa unidade lexical ocorre como base nominal de uma relação de coocorrência.

Isso tem um impacto importante na aprendizagem da língua e no domínio de mecanismos de codificação, tanto pela pertinência das informações explicitamente fornecidas nos lugares lexicográficos em que se espera e se busca tal informação (*face visível*) quanto pela exposição acidental, mas sempre possível, às diferentes possibilidades de realização sintagmática desses mecanismos em diferentes espaços topológicos do discurso lexicográfico.

O mesmo pode ser observado em *prestige: [...] avoir du prestige*, cujo enunciado lexicográfico apresenta um exemplo (*avoir du prestige*) que, enquanto informação sintático-semântica explicitamente forne-

cida, corresponde a uma *face visível* da palavra que, por sua vez, se realiza e se legitima tanto na descrição de *prestigieux*, *euse: qui a du prestige* quanto em enunciados definicionais dispersos no discurso lexicográfico como, por exemplo, em *influent*, *ente: qui a du prestige*.

Esse tipo de circularidade entre diferentes enunciados lexicográficos, representada pela presença de uma mesma informação sintático-semântica que transita entre eles, pode ser facilmente observada em **a.** *prudence: [...] avoir de la prudence ↔ prudent, ente: qui a de la prudence*; **b.** *influence [...] avoir de l'influence ↔ influent, ente: qui a de l'influence*; **c.** *préjudice: [...] porter préjudice ↔ fâcheux, euse: qui porte préjudice*; **d.** *jalousie [...] éprouver de la jalousie ↔ jaloux, ouse: qui éprouve de la jalousie*; **e.** *dépit: [...] éprouver du dépit ↔ dépité, ée: qui éprouve du dépit*, assim como em **f.** *importance: [...] avoir de l'importance ↔ importer: avoir de l'importance ↔ important, ante: qui a de l'importance ↔ intéresser: avoir de l'importance* etc.

É possível observar que essas relações implícitas que se estruturam em torno de diferentes enunciados lexicográficos são de extensão variável e podem formar redes mais extensas de interface entre eles, como acontece em **a.** *connaissance: [...] perdre connaissance ↔ (s') évanouir: perdre connaissance ↔ évanouissement: le fait de perdre connaissance ↔ inanimé, ée: qui a perdu connaissance ↔ noyé, ée: qui a perdu connaissance ↔ perdre: [...] perdre connaissance*, assim como em **b.** *plaisir: [...] prendre plaisir à ↔ cruel, cruelle: qui prend plaisir à ↔ (se) plaire: prendre plaisir à ↔ sadique: qui prend plaisir à ↔ taquin: qui prend plaisir à* e em **c.** *atteinte: [...] porter atteinte à ↔ attentatoire: qui porte atteinte à ↔ déshonorer: porter atteinte à ↔ diffamation: qui porte atteinte à ↔ diffamatoire: qui tend à porter atteinte à ↔ diffamer: chercher à porter atteinte à ↔ discréditer: porter atteinte à ↔ violer: porter atteinte à ↔ fraude: ...l'intention de porter atteinte à*. Neles podemos observar claramente que a descrição da base nominal (*connaissance/*

plaisir/atteinte) comum a vários enunciados definicionais fornece explicitamente (*face visível*) informações sintático-semânticas sobre uma combinação lexical na qual não há intervenção de determinantes entre coocorrentes (*perdre connaissance/prendre plaisir/porter atteinte*), apresentando e legitimando mecanismos sintático-semânticos que podem ser observados de forma dispersa no discurso lexicográfico.

Dentro desse quadro de circularidade, é interessante observar, no entanto, que a *face visível* da palavra nem sempre oferece as melhores condições para que o usuário possa observar as coerções sintático-semânticas que se impõem às relações de coocorrência de uma unidade lexical. Entre **a.** *mépris*: [...] *n'avoir que du mépris* e **b.** *méprisant*: *qui a du mépris*, assim como entre **c.** *dédain*: [...] *n'avoir que du dédain* e **d.** *dédaigneux*: *qui a du dédain*, é possível observar que, diferentemente do que acontece nos enunciados definicionais que informam claramente sobre as condições de uso desses pares de coocorrentes (*qui a du mépris/qui a du dédain*), a intervenção de suportes gramaticais complexos (*ne...que*) na descrição da base lexical (*n'avoir que du mépris/n'avoir que du dédain*) compromete a visibilidade de mecanismos importantes para o desenvolvimento da capacidade de codificação em língua francesa e, por conseguinte, a observação e reconhecimento de informações que estruturam relações de coesão dessa natureza no interior do discurso lexicográfico.

O caso a seguir trata de dois pares de coocorrentes (1. *faire mal, faire du mal*; 2. *avoir mal, avoir du mal*) que aparecem *espontaneamente* e de forma dispersa em diferentes enunciados lexicográficos. É possível observar claramente que cada um desses pares se estrutura por diferenças sintático-semânticas que opõem a presença ou a ausência de determinantes e, por conseguinte, diferenças de sentido importantes. Trata-se, naturalmente, de casos de coocorrência que podem causar dificuldade para quem aprende francês, sobretudo porque eles

se estruturam, de forma cruzada, por relações de semelhança (1. *faire du mal, avoir du mal*; 2. *faire mal, avoir mal*) e, de forma paralela, por relações de dissemelhança (1. *faire mal, faire du mal*; 2. *avoir mal, avoir du mal*).

Na descrição de *mal*, é possível observar a presença de exemplos que informam sobre as coerções sintático-semânticas que se impõem ao uso dessa unidade lexical nesses dois pares de coocorrência. Isso torna esses exemplos o ponto de articulação principal de um circuito fechado e coeso de informações constituído pelo verbete (*mal*) que informa sobre essas configurações coocorrenciais e todos aqueles que apresentam ocorrências *accidentais* desses pares de coocorrentes no interior do discurso lexicográfico como um todo.

Em *mal*, podemos observar um princípio de codificação formal que apresenta as informações com maiúsculas e negrito, a fim de dar mais visibilidade à *face visível* da palavra: *mal, maux: I. 2. □ AVOIR MAL: souffrir □ FAIRE MAL: faire souffrir*. Esse procedimento tem efeito reforçado pela proximidade dessas duas informações, o que se torna possível por se tratar de uma mesma aceção do substantivo. Contrastivamente, na apresentação de *faire du mal/avoir du mal*, uma *ruptura* desse código de estruturação formal (uso de fonte minúscula, sem negrito) pode ser observada, no mesmo verbete: *mal, maux: I.1. [...] faire du mal à qqn [...] 5. [...] avoir du mal à faire qqch*, comprometendo a percepção de uma *cumplicidade* entre essas informações, assim como a visibilidade de microsistemas de relações de oposição/semelhança sintático-semântica entre elas, conforme foi comentado anteriormente. Isso se agrava pelo distanciamento, no espaço topológico do verbete, de uma parte dos elementos desse *jogo* de oposições e semelhanças, por se tratar de aceções dessa vez diferentes do substantivo.

Na verdade, são muito variadas as condições de visibilidade dadas pelo enunciado lexicográfico para que o usuário observe informações importantes sobre o funcionamento da língua, a fim de que possa produzir enunciados de acordo com as exigências sintático-semânticas próprias do idioma. Isso significa dizer que as informações explicitamente fornecidas sobre uma unidade lexical são, naturalmente, um aspecto da *face visível* da palavra que, no entanto, pode apresentar informações de forma mais ou menos velada, de acordo com as condições de visibilidade dadas. Esse aspecto da descrição lexicográfica merece um estudo mais detalhado, por ter muito a contribuir para o aperfeiçoamento da vocação didática dos dicionários, sobretudo pedagógicos.

2.2. A *face invisível* das palavras

Como já foi dito, estamos entendendo como *face invisível* da palavra as ocorrências de uma unidade lexical em enunciados definicionais dispersos no discurso lexicográfico, nos quais se encontram informações sintático-semânticas sobre a relação dessa unidade com seus coocorrentes privilegiados, sem que, no entanto, elas sejam fornecidas na própria descrição dessa unidade lexical.

A *face invisível* da palavra pode ser observada em enunciados definicionais, como é o caso de **a.** *bienveillant: qui a de la bienveillance*, **b.** *malveillant: qui a de la malveillance*, **c.** *capable: qui a de l'habileté*, **d.** *adroit: qui a de l'habileté*, **e.** *industrieux: qui a de l'habileté*, **f.** *industrieux: qui a de l'adresse* e **g.** *adroit: qui a de l'adresse*, quando percebemos que, na descrição dessas unidades lexicais (*bienveillance*, *malveillance*, *habileté*, *adresse*), não há exemplos que informem, respectivamente, sobre as coocorrências *avoir + bienveillance*, *avoir + malveillance*, *avoir + habileté*, *avoir + adresse*, de modo que o usuário seja informado sobre as coer-

ções sintático-semânticas que se impõem a essas combinações lexicais, cuja relação é intermediada pela intervenção de determinantes partitivos (*avoir de la bienveillance, avoir de la malveillance, avoir de l'habileté, avoir de l'adresse*). Considerando que um usuário estrangeiro que aprende a língua não tem condições de predizer quais mecanismos estruturam essas relações, a *face invisível* das palavras relega o usuário à descoberta dessas informações por meio de encontros fortuitos com essas formas em enunciados lexicográficos dispersos no dicionário.

Também é interessante notar que diferentes formas de coocorrência de uma unidade lexical podem ser oferecidas de forma pulverizada em diferentes enunciados lexicográficos, de modo que as diferentes *faces invisíveis* das palavras só podem se revelar ao usuário em episódios fortuitos de consulta, os quais lhe reservam a oportunidade de descobrir *acidentalmente* fisionomias da palavra que não foram oferecidas e não puderam ser encontradas no próprio enunciado lexicográfico em que elas poderiam figurar. Isso pode ser observado em **a. dégoûtant: qui inspire du dégoût**, **b. dégoûté: qui éprouve du dégoût**, **c. dégoûter: donner du dégoût** e **d. déplaire: causer du dégoût**, que oferecem um leque de formas de coocorrência que, enquanto faces invisíveis da palavra, relegam essas descobertas a encontros casuais no discurso lexicográfico, em que a consulta já não é mais motivada pela mesma necessidade de informação. Nesse caso, em particular, a pulverização dessas informações no interior do discurso lexicográfico contrasta com a indigência da *face visível* da palavra, na qual se observa apenas uma coocorrência da mesma natureza: *dégoût: [...] avoir du dégoût*. Regidas por coerções sintático-semânticas análogas, essas *faces invisíveis* da palavra poderiam contribuir para desenvolver no usuário uma intuição quanto às possibilidades de intervenção dessas coerções, caso elas fossem diretamente observáveis por ele.

Em **a.** *complaisant*, ante: *qui a de la complaisance* e **b.** *courageux*, *euse*: *qui a du courage*, a *face invisível* que o enunciado definicional oferece de *complaisance* e de *courage* (*avoir de la complaisance* e *avoir du courage*) pode ser explicada pela ausência dessa informação na descrição dessas unidades lexicais. Na descrição de *complaisance* e *courage*, os exemplos fornecidos explicitamente (*auriez-vous la complaisance de m'ouvrir la porte* e *je n'ai pas le courage de me lever si tôt*, respectivamente) dão, como *face visível* da palavra, informações sintático-semânticas (*avoir la complaisance de* e *avoir le courage de*, respectivamente) que, na verdade, derivam da configuração sintagmática coocorrential de *avoir de la complaisance* e *avoir du courage*, em função da intervenção de operações de qualificação que introduzem complementos nominais, uma vez que **avoir de la complaisance de* e **avoir du courage de* não são possibilidades da língua. Ainda que haja uma relação genética entre essas informações, essa face da palavra (*avoir de la complaisance/avoir du courage*) permanece invisível porque ela não é predizível, para um usuário estrangeiro que aprende a língua e consulta o dicionário, a partir dos mecanismos explicitamente dados.

Em **a.** *intéresser*: I. 2. *avoir de l'intérêt* e **b.** *valoir*: I. A. 3. *avoir de l'intérêt*, podemos considerar que a informação sobre a coocorrência *avoir + intérêt* e as condições sintático-semânticas em que ela acontece (*avoir de l'intérêt/*avoir intérêt*) fornece uma face oculta da palavra quando observamos que essa informação não é dada em *intérêt*, ainda que essa forma de coocorrência seja utilizada em diferentes enunciados definicionais e seja comum a acepções diferentes dessa unidade lexical: 1. ter importância (acepção 3), na qual *avoir de l'intérêt* permitiria a produção de frases como *cette étude a de l'intérêt pour toute la communauté scientifique* e 2. interessar-se por algo (acepção 6), na qual *avoir de l'intérêt* permitiria a produção de frases como *j'ai de l'intérêt pour ce genre de musique*, que são exemplos dados aqui aleatoriamente.

Por outro lado, para a compreensão de oposições sintático-semânticas importantes na língua, a visibilidade dessa informação daria condições para que, na codificação de enunciados, o usuário pudesse observar e distinguir o que opõe *avoir de l'intérêt* e *avoir intérêt*, que, contrariamente à primeira, é explicitamente fornecida na descrição de *intérêt*.

Em **a.** *absent: qui ne porte pas attention à ce qui l'entoure* e **b.** *étourdi: qui ne porte pas attention à ce qu'il fait*, podemos observar que os enunciados definicionais revelam uma *face invisível* da palavra, à medida que verificamos que, na descrição de *attention*, não há nenhum exemplo que informe sobre a relação de coocorrência dessa unidade lexical com *porter* e, por conseguinte, sobre as coerções sintático-semânticas que, por um lado, estruturam essa relação (ausência de determinante) e que, por outro, subsumem aquelas que se impõem à construção negativa (ausência de partícula de determinante), que, nesse caso, se apresenta *acidentalmente* em enunciados definicionais (*absent/étourdi*) dispersos no discurso lexicográfico. Em termos lexicográficos, é importante considerar, naturalmente, a inviabilidade de fornecer exaustivamente os coocorrentes possíveis de uma unidade lexical e, por conseguinte, de informar sobre as condições sintático-semânticas que sustentam a relação entre eles. Nessas condições, ainda que não seja ideal, a *face invisível* da palavra é uma chave de acesso a informações sobre a língua que dará ao usuário o prazer de descobertas que sempre se insinuam no discurso lexicográfico, dando um significado cada vez maior ao que se chama de riqueza dos dicionários.

Em **a.** *intéresser: prendre intérêt*, **b.** *soucier: prendre intérêt* e **c.** *ennui: fait qu'on ne prend d'intérêt à rien*, os três enunciados definicionais oferecem uma *face invisível* da palavra ao informar sobre uma relação de coocorrência (*prendre + intérêt*) que não é fornecida na descrição de

intérêt. Essa informação encontra-se, portanto, dispersa em diferentes enunciados lexicográficos. É interessante notar, no entanto, que há um conflito de informações entre eles, à medida que o terceiro (*ennui: fait qu'on ne prend d'intérêt à rien*) informa, por meio de frase negativa, sobre coerções sintático-semânticas que não são compatíveis com aquelas dos primeiros enunciados definicionais (*intéresser: prendre intérêt* e *soucier: prendre intérêt*), uma vez que a presença de uma partícula negativa (*d'*) no terceiro enunciado definicional pressupõe a intervenção de determinantes na estruturação dessa relação de coocorrência (*prendre de l'intérêt*), o que não se observa nos dois primeiros, nos quais há ocorrência de determinante zero.

Em outro caso semelhante, é difícil decidir se há conflito de informações no interior do discurso lexicográfico ou concorrência entre duas formas equivalentes de coocorrência (*porter de l'intérêt/ porter intérêt*). Em *se désintéresser: ne plus porter intérêt à*, o enunciado definicional oferece informações sintático-semânticas que se observam claramente na forma negativa (ausência de partícula de determinante), legitimando as coerções sintagmáticas que subjazem a esse enunciado (*porter intérêt à*). No entanto, em *intérêt: [...]* *porter de l'intérêt à*, o que se observa, no exemplo, é uma *face visível* da palavra que informa sobre uma coocorrência na qual intervêm condições sintático-semânticas diferentes daquelas que se observam no enunciado definicional *disperso* no discurso lexicográfico.

A *face invisível* da palavra nem sempre se deixa observar facilmente no discurso lexicográfico, pois ela pode se apresentar sob diferentes formas de configuração sintagmática, em função das possibilidades sintático-semânticas que derivam das coerções que dão origem a essas possibilidades. Vimos rapidamente que a frase negativa e operações de qualificação, por exemplo, obedecem a coerções que se

apresentam como possibilidades derivadas daquelas que regem diferentes relações de coocorrência.

Em *improbable: qui a peu de chances de se produire*, a operação de quantificação (*peu de*) dissimula um pouco as coerções sintático-semânticas que se observam melhor em *outsider: qui a des chances de gagner*. Esses enunciados definicionais permitem observar a predominância de determinante plural em uma acepção particular de *chance* (3. *possibilité de se produire par hasard*). Eles revelam uma *face invisível* da palavra à medida que informam sobre uma forma de coocorrência (*avoir des chances*) que não é fornecida na descrição de *chance*, apesar da importância que teria fornecer essa informação como *face visível* da palavra, pois isso poderia permitir a apreensão de um jogo importante de oposições sintático-semânticas entre *avoir des chances* e *avoir de la chance*. É interessante notar que, diferentemente da primeira forma de coocorrência, a segunda é fornecida explicitamente pela descrição: *chance: 3. [...] avoir de la chance, beaucoup de chance*. Nesse último caso, vamos observar uma coesão entre verbetes, uma vez que a *face visível* da palavra autoriza e legitima fórmulas definicionais dispersas no dicionário: **a.** *chançard: qui a de la chance*, **b.** *chanceux: qui a de la chance* e **c.** *privilegié: qui a de la chance*. Em ambos os casos entram em jogo os mesmos coocorrentes, submetidos, no entanto, a coerções sintático-semânticas diferentes.

Para concluir brevemente, é importante dizer que, apesar do valor didático de uma apresentação formal que permita observar e confrontar informações (*face visível*), a fim de permitir a percepção de aspectos importantes na língua, a complexidade do trabalho lexicográfico e a dificuldade de controle das informações no interior do discurso lexicográfico parecem gerar inevitavelmente situações em que algumas descobertas sobre a língua permaneçam por conta do feliz acaso e da capacidade observadora do usuário.

CONSIDERAÇÕES FINAIS

Podemos considerar a *face invisível* da palavra como um aspecto inerente do discurso lexicográfico, dada a diversidade e complexidade de tanto das informações que o constituem quanto das relações que implicitamente se estabelecem entre diferentes enunciados lexicográficos, formando redes de informação que, muitas vezes, não são diretamente observáveis nem pelo usuário nem pelo próprio lexicógrafo. Um estudo minucioso de diferentes aspectos da *face invisível* da palavra encontra seu verdadeiro sentido em contraposição com sua face complementar, visível e observável no enunciado lexicográfico que, dentre todos os que podem revelar faces invisíveis da palavra, corresponde ao lugar privilegiado e por excelência para informar sobre uma unidade lexical.

É importante observar que a *face invisível* da palavra não aponta necessariamente para insuficiências na descrição lexicográfica, pois o discurso lexicográfico sempre reservará ao usuário espaços permanentes de descobertas sobre a língua. Por outro lado, no entanto, a contraposição de duas faces pode contribuir para melhorar aspectos da descrição lexicográfica (*face visível*) que são importantes para a qualidade técnica da descrição e para promover a vocação didática de uma obra lexicográfica, facilitando assim a aprendizagem de uma língua por meio dessa ferramenta tão importante, que é o dicionário.

REFERÊNCIAS

CALDERÓN CAMPOS, Miguel. **Sobre la elaboración de diccionarios monolingües de producción**. Granada: Universidad de Granada, 1994.

FRADIN, Bernard; MARANDIN, Jean-Marie. Autour de la définition: de la lexicographie à la sémantique. **Langue française**, Paris, n. 43, p. 60-83, 1979.

HWANG, Alvaro David. Dicionário Pedagógico: a sistematicidade como princípio de redundância pedagógica. *In*: NADIN, Odair; RODRIGUES-PEREIRA, Renato. (org.). **Lexicografia pedagógica**: pesquisas, práticas e perspectivas. Campinas: Mercado de Letras, 2023. p. 67-94.

LE PETIT ROBERT. **Dictionnaire alphabétique et analogique de la langue française 2011**. Version 3.4. Paris: Dictionnaires Le Robert, c2010. CD-ROM.

MARTINEZ, Camille; PRUVOST, Jean. De la traque lexicographique du sens et de la triple investigation dictionnaire. **Synergies Tunisie**, Gerflint, p. 139-151, 2010. Disponível em: <https://gerflint.fr/Base/Tunisie2/pruvost.pdf>. Acesso em: 21 jun. 2021.

MEDEROS, Humberto. A propósito de la definición lexicográfica. *In*: HERNANDEZ, Humberto *et al.* (org.). **Aspectos de lexicografía contemporánea**. Barcelona: Biblograf, 1994, p. 95-106.

PRUVOST, Jean. Quelques concepts lexicographiques opératoires à promouvoir au seuil du XXI^e siècle. **Études de linguistique appliquée**, v. 37, n. 1, p. 7-37, 2005. Disponível em: <https://www.cairn.info/revue-ela-2005-1-page-7.htm>. Acesso em: 21 jun. 2021.

QUEMADA, Bernard. Note sur lexicographie et dictionnaire. **Cahiers de Lexicologie**, Paris, v. 51, n. 2, p. 235-242, 1987.

REY, Alain. Lexicographie. Table ronde. **Lexiques**. Paris: Le français dans le monde, 1989.

REY-DEBOVE, Josette. Étude **linguistique et sémiotique des dictionnaires français contemporains**. The Hague/Paris: Mouton, 1971.

REY-DEBOVE, Josette. Statut et fonction de l'exemple dans l'économie du dictionnaire. *In*: HEINZ, Michaela. **L'exemple lexicographique dans les dictionnaires français contemporains**. Actes des premières journées allemandes des dictionnaires. Tübingen: Max Niemeyer, 2005. p. 15-20.

GÍRIA: MARCA DE USO PRESENTE EM DICIONÁRIOS ESCOLARES

Claudia Zavaglia

Fábio Henrique de Carvalho Bertonha

INTRODUÇÃO

Nesta pesquisa, interessa-nos discorrer sobre as marcas de uso, restringindo-nos à marca diastrática ‘gíria’. Entendemos que o estudo se baseia na presença e frequência do uso de gírias por determinada comunidade de falantes, principalmente jovens, sendo que sua consagração nos dicionários se dá pelo tempo por meio do conhecimento oral dos povos, possibilitando sua cristalização.

Sob tal perspectiva, objetivamos uma recolha da marca de uso ‘gíria’ presente em oito dicionários monolíngues, o *corpus* desta pesquisa, a saber: Caldas Aulete (Geiger, 2011), Dicionário Aurélio ilustrado (Ferreira, 2008), Fala Brasil! (Braga; Magalhães, 2011), Dicionário ilustrado de português (Biderman, 2009), Novo Dicionário Eletrônico Aurélio (Ferreira, 2010), Dicionário Eletrônico Houaiss da Língua Portuguesa (Houaiss, 2009), Dicionário Houaiss da Língua Portuguesa (Houaiss, 2021) e Aulete Digital (Aulete, 2021), respectivamente, doravante CA, DAI, Fala Brasil!, DIP, AU, HO, HOvir, CAvir. Pretendemos analisar a dicionarização de sentidos marcados por essa etiqueta em algumas entradas lexicográficas e,

por conseguinte, sua efetiva contribuição (ou não) na restrição de significados e usos.

Almejamos, pois, a partir de verbetes que contêm a marca ‘gíria’, analisar as seguintes questões: (i) as especificidades dos dicionários que tomamos como fonte de pesquisa; (ii) a perspectiva pragmática e variação linguística no uso dos itens lexicais, por exemplo, de ordem social, de registro e também de limites geográficos, nos dicionários; e (iii) a complexidade do fenômeno da gíria na língua.

1. PERCURSOS TEÓRICOS

Iniciamos por refletir sobre alguns aspectos lexicográficos relacionados aos tipos de dicionários do *corpus* com a intenção de analisarmos as especificidades de seus verbetes, sobretudo em relação às informações microestruturais que compõem os verbetes dessas obras, com especial atenção às marcas de uso. Para tanto, tomamos por base os pressupostos teóricos de Porto Dapena (2002), Welker (2008), Pontes (2009), Novoa (2005), Krieger (2006) e Vilarinho (2013). Na sequência, abordaremos a perspectiva pragmática pela variação linguística e sua relação com a gíria pela contribuição dos estudos de Preti (2000), Bagno (2007), Patriota (2009), Coelho *et al.* (2010; 2015).

1.1. Dicionários escolares e gerais

Considerando-se os dicionários como obras de consulta que se prestam a descrever o léxico, Porto Dapena (2002, p. 42-43) afirma que sua tipologia pode ser estabelecida a partir de quatro variáveis: (i) número e extensão das entradas; (ii) modo de estudar as entradas; (iii) ordenação utilizada; e (iv) suporte da obra. À vista disso, um dicionário geral almeja, utopicamente, abarcar a totalidade do léxico, uma vez

que a língua é dinâmica; entretanto, não é possível contemplar ‘todo’ o repertório lexical de uma língua descrevendo, em verbetes, as informações morfosintáticas, semântico-pragmáticas de um item lexical.

Já as obras dicionarísticas escolares podem ser vistas como reduções dos dicionários gerais, pois são um recorte da língua para a produção direcionada a um público-alvo, de acordo com sua mudança e avanço educacionais (Vilarinho, 2013, p. 62). Em vista disso, a nomenclatura desse tipo de dicionário visa a selecionar unidades lexicográficas que abarquem um subconjunto do léxico voltado a estudantes em fase escolar. Assim, a proposta do Programa Nacional do Livro Didático (PNLD), ramificando-se no PNLD 2012: Dicionários (Brasil, 2012) divide-se em quatro categorias em razão do consulente, sua escolaridade, além do volume informacional inserido nos dicionários, conforme retrata o Quadro 1 a seguir:

Quadro 1 – Tipologias estabelecidas pelo PNLD 2012: Dicionário

Tipo de dicionário	Fase educacional em que o estudante se encontra	Características
Tipo 1	1º ano do EF	Proposta lexicográfica apropriada às demandas do processo de alfabetização inicial; de 500 a 1.000 verbetes.
Tipo 2	Período entre o 2º e o 5º ano do EF	Proposta lexicográfica apropriada a estudantes em fase de consolidação do domínio tanto da escrita quanto da organização e da linguagem típicas do gênero dicionário; de 3.000 a 15.000 verbetes.
Tipo 3	Segundo segmento do EF (6º ao 9º ano)	Proposta lexicográfica direcionada pela caracterização de um dicionário padrão de uso escolar, porém, apropriada aos estudantes que se encontram nos últimos anos do Ensino Fundamental; de 19.000 a 35.000 verbetes.
Tipo 4	1º ao 3º ano do EM	Proposta lexicográfica própria de um dicionário padrão, porém, apropriada às demandas escolares do Ensino Médio, inclusive o profissionalizante; de 40.000 a 100.000 verbetes.

Fonte: Brasil (2012)

Concernente à área que deva se ocupar do tratamento lexical estudantil, Krieger (2006, p. 247) destaca que o grupo de aspectos envolvendo os dicionários no ambiente escolar passou a ser entendido como Lexicografia Didática, subdivisão linguística que se ocupa dos níveis educacionais e, assim,

[...] como uma produção direcionada à escola é de extrema importância, sobretudo porque há uma tendência geral de identificar como escolar os dicionários tipo mini. No entanto, a compreensão do caráter escolar costuma estar associada mais às suas dimensões reduzidas do que à sua efetiva adequação ao ensino/aprendizagem da língua. Por isto, apesar de práticas, as versões sintéticas nem sempre são as melhores para uso escolar.

Voltando-se ao público escolar, busca-se repertoriar os itens lexicais particulares ao universo educacional de seus consulentes, obviamente, oferecendo-se um juízo crítico avaliativo expresso pelos lexicógrafos com relação à Lexicografia Escolar. Não vamos nos aprofundar nas discussões sobre como nomear essa área de pesquisa, mas é importante destacar que não há consenso entre seus estudiosos. Assim, ainda para a caracterização das obras dicionarísticas escolares, Welker (2008, p. 295) ressalta que “fazem parte dos dicionários pedagógicos de língua materna (DPLMs) não somente os dicionários escolares como também, por um lado, os dicionários infantis (DIs)”.

Notamos também que não há uniformidade para se tratar dessa área lexicográfica, visto que tanto Climent de Benito (2005) quanto Corbin (2001) afirmam que esses tipos de dicionários, na verdade, correspondem, respectivamente, a um Dicionário de Iniciação e a um Pré-dicionário (Welker, 2008). Já Biderman (1984) considera Dicionário Escolar como Dicionário Infantil. Embora haja divergência terminológica, Pontes (2009) destaca que as obras dicionarísticas

apresentam uma natureza didática, uma vez que comunicam conhecimentos linguísticos (informações fonéticas, morfofossintáticas, semântico-pragmáticas) a seus consulentes, configurando-se em excelentes fontes pedagógicas de ensino.

Nota-se que o não estabelecimento de limites claros entre obras dicionarísticas escolares e infantis contribui para um não reconhecimento imediato do conteúdo nem do público-alvo, haja vista a inexistência terminológica coincidente, sendo encontrados Dicionário Escolar, Dicionário Infantil, Dicionário Escolar Infantil, Dicionário de Iniciação, Pré-dicionários. Para atender ao escopo desta pesquisa, utilizamos a concepção de obra lexicográfica escolar empregada pelo PNLD 2012: Dicionários, na qual o dicionário é um tipo de obra que

[...] não só se mostra compatível com essas atividades como é pensada para propiciar o seu desenvolvimento; e, entre eles [os dicionários], são ainda mais adequados os que foram concebidos e elaborados para atender a essas demandas específicas. Como uma dessas demandas [de ensino e aprendizagem] é exatamente a da adaptação do que se quer ensinar/aprender ao nível de ensino e aprendizagem visados, podemos acreditar que os dicionários orientados para faixas específicas serão mais eficazes em seus propósitos pedagógicos. Na medida em que os dicionários escolares disponíveis no mercado livreiro visam diferentes públicos, obedecem a diferentes propostas e são realizados com graus variados de rigor, podem se revelar mais ou menos adequados para a consecução dos objetivos pedagógicos visados (Brasil, 2012, p. 18).

Ao se elaborar uma obra dicionarística voltada ao público escolar, é fundamental ter em vista a finalidade de alcançar esse público, inserindo-o no fascinante mundo das palavras. Portanto, é necessário

que seja pensada a formatação da obra para que os estudantes possam se familiarizar cada vez mais com a estrutura do dicionário infantil. Apesar de se tratar de uma obra de dimensões reduzidas, nos dicionários escolares, é possível encontrar certo acervo que corresponde à herança linguístico-histórico-cultural do país.

Por fim, durante as fases de alfabetização e letramento, o repertório vocabular dos estudantes, cada vez mais, vai se ampliando à medida que seus conhecimentos vão aumentando; por isso, a cada fase de sua vida, determinadas palavras são mais usadas do que outras posto que são, de fato, um reflexo da inserção social de cada estudante. Logo, é necessário pensar em dicionários voltados ao público escolar que possam contribuir para expansão vocabular do indivíduo e o seu melhor desempenho na sua língua materna no decorrer desse percurso estudantil.

1.2. Variação linguística e a marca de uso ‘gíria’

No momento em que se busca um sentido nos dicionários, é preciso interpretá-lo (con)textualmente, com vistas às várias áreas do conhecimento, além dos campos semânticos, que compõem essa descrição informacional. E uma das formas de se restringir e se valorar a diversidade significativa de um item lexical é, justamente, a partir da inserção de marcas de uso na microestrutura das entradas das obras lexicográficas, cuja compreensão, não raro, é alcançada intuitivamente pelos consulentes, e, por vezes, de modo arbitrário.

No que tange à variação linguística, seu entendimento advém da compreensão da língua como heterogênea, conforme afirmam Coelho *et al.* (2010, p. 20), fundamentados em postulados labovianos. Didaticamente, Bagno (2007, p. 36) explica o conceito de heterogeneidade da língua:

Ao contrário da norma-padrão, que é tradicionalmente concebida como um produto homogêneo, como um jogo de armar em que todas as peças se encaixam perfeitamente umas nas outras, sem faltar nenhuma, a língua, na concepção dos sociolinguistas é intrinsecamente heterogênea, múltipla, variável, instável e está sempre em desconstrução e reconstrução. Ao contrário de um produto pronto e acabado, de um monumento histórico feito de pedra e cimento, a língua é um processo, um fazer-se permanentemente e nunca concluído. A língua é uma atividade social, um trabalho coletivo, empreendido por todos os seus falantes, cada vez que eles se põem a interagir por meio da fala ou da escrita.

Sob a perspectiva sociolinguística, as línguas sofrem variações, logo, são heterogêneas. A partir disso, para Coelho *et al.* (2015, p. 23), quando um indivíduo faz uso do item lexical, conseqüentemente, evidencia que a variação pode ser sistematizada, quer dizer, não se trata de uma balbúrdia linguística. Em consonância, para Bagno (2007, p. 40), a variação não é desorganizada e caótica na língua, longe disso: as mudanças e variações ocorrem ordenadamente, condicionadas a variados fatores.

Outro ponto a ser trazido à baila é a afirmação de que o brasileiro não gosta de sua língua por considerá-la difícil e que, portanto, ocorreriam conflitos entre a língua padrão e suas variantes. Acerca dessa afirmação Bagno (2007, p. 35) destaca que,

[...] como o nosso ensino da língua sempre se baseou na norma gramatical de Portugal, as regras que aprendemos na escola em boa parte não correspondem à língua que realmente falamos e escrevemos no Brasil. Por isso achamos que “português é uma língua difícil”: porque temos de decorar conceitos

e fixar regras que não significam nada para nós. No dia em que nosso ensino de português se concentrar no uso real, vivo e verdadeiro da língua portuguesa do Brasil é bem provável que ninguém mais continue a repetir essa bobagem.

Esse purismo linguístico, que remonta às origens do Brasil, sobretudo com a vinda da corte de Portugal, foge à própria formação cultural heterogênea, gerando, assim, um ambiente preconceituoso contra aqueles indivíduos que possuem lacunas de conhecimento da gramática normativa. Patriota (2009) ressalta que é a opção de uso escolhido como o “melhor”, em detrimento dos outros, que será o padrão a ser seguido, sendo que aquilo que estiver fora dessa régua normativa é considerado marginal.

Observando que há constantes variações na língua, nos dicionários, é preciso que seu registro aconteça e, para que sua dicionarização contribua para esclarecer e/ou ampliar o conhecimento dos consultantes, as marcas de uso podem desempenhar papel essencial. Nesta pesquisa, temos como foco as marcas diastráticas, ou seja, aquelas que sinalizam os usos dos falantes ligados a determinados grupos sociais, que destacam os itens lexicais utilizados por determinadas parcelas da sociedade. Novoa (2005, p. 33) aponta que

[...] através dessas marcas, as restrições de uso são indicadas em termos de raça ou etnia, diferenças de idade, diferenças sexuais, classe socioeconômica, profissões, nível educacional, formação cultural, etc. Além disso, as indicações abundam no jargão profissional, gíria do mundo dos delinquentes etc.

Referindo-nos às propostas de Hausmann (1977 *apud* Welker, 2004), as marcações diastráticas são assim caracterizadas: ‘elevado’,

‘coloquial’, ‘chulo’, ‘gíria’, ‘familiar’, ‘popular’, ‘vulgar’, ‘tabu’. Como aponta Novoa (2005), existe uma linha tênue entre as marcas diastráticas e as diafásicas, quer dizer, aquelas cujo estilo de linguagem se registra como formal ou informal (variação estilística); por isso, nesta pesquisa, iremos focar apenas na marca diastrática ‘gíria’, uma vez que essa variação social se caracteriza pelo sentimento de pertencimento, criando uma identidade de grupo a partir de suas características próprias que alcançam o léxico.

Por fim, a gíria é uma variante linguística presente e eventualmente, marcada nos dicionários. Segundo Preti (2000, p. 64), esse item lexical, geralmente, apresenta-se em situações de comunicação em que os falantes são menos cultos ou em diálogos menos formais, por isso essas lexias gírias, com frequência, são etiquetadas com outras marcas, tais como ‘vulgar’, ‘familiar’, ‘popular’ etc. Essa não padronização das marcas evidencia um tratamento individual do lexicógrafo, podendo ser motivado por estereótipos, associando a gíria à marginalidade.

2. PERCURSOS METODOLÓGICOS

A partir do *corpus* de dicionários selecionado, coletamos verbetes que continham a marca de uso ‘gíria’ em suas microestruturas:

Quadro 2 – ‘Gíria’ em dicionários do Tipo 2

Dicionário	Volume de verbetes encontrados	Unidades lexicográficas marcadas por ‘gíria’
CA (2011)	78	abacaxi, abafado, abobrinha, agito, alcaguete, alegre, apelar, apitar, arrebentar, avacalhar, azaração, azarar, bacana, bafo, balada, bandeira, barato, bico, biruta, boca, bode, bola, bolha, borrar, brasa, bruxo(a), camelo, cano, cantar, canudo, capotar, cara, careta, cascata, chulé, chute, colar, colher, conversa, coroa, dançar, dedo-duro, duro, espírito, fim, furar, galera, galho, invocado, irado, manjar, mina, mole, moleza, morar, onda, pau, pegar, pescar, pintar, pivete, pronto, rolar, secar, sinistro, sopa, sujar, trabalhar, trouxa, tutu, vacilar, valer, viajar, xará, zoar, zona, zonear, zorra
DAI (2008)	28	abafar, abobrinha, alugar, barbeiro, bico, bife, biruta, bobear, chiar, cuca, dureza, duro, grana, irado, legal, limpo, matar, mofar, morar, paquerar, pau, pré-histórico, pronto, quicar, tira, treco, troço, zoar
Fala Brasil! (2011)	-----	-----
DIP (2009)	-----	-----

Fonte: elaborado pelos autores

As obras do Quadro 2 fazem parte da seleção e categorização realizadas pelo PNL D 2012: Dicionários (Brasil, 2012) e correspondem ao período educacional entre o 2º e o 5º ano do EF. Esses quatro dicionários se caracterizam por terem um mínimo de 3.000 e um máximo de 15.000 verbetes, cuja proposta lexicográfica se adequa a alunos que estejam em fase de consolidação do domínio tanto do sistema gráfico da escrita quanto da organização e da linguagem comuns ao gênero dicionário. Como se observa, dois desses dicionários negligenciam o uso da marca ‘gíria’ em seus verbetes.

Dentre as entradas coletadas, elencamos nove delas: *abobrinha*, *bico*, *biruta*, *duro*, *irado*, *morar*, *pau*, *pronto* e *zoar*. Em seguida, essas unidades lexicográficas foram postas em confronto com suas homônimas em quatro dicionários gerais (dois impressos e dois *on-line*) com o escopo de verificarmos se as restrições de uso empregadas se mantêm com as mesmas ressalvas de uso contextual nas obras dedicadas aos estudantes. Resulta que, dos oito dicionários selecionados, dois escolares (Fala Brasil! e DIP) não apresentaram a marca ‘gíria’, portanto, apenas CA e DAI foram contrastados com os quatro gerais, utilizados como comparativos para balizar nossas análises.

3. ANÁLISE E RESULTADOS ALCANÇADOS

Nossas análises se iniciam a partir de nove entradas que serão reproduzidas e adaptadas aqui, que possuem a marcação ‘gíria’ em suas microestruturas, conforme explanado na seção anterior.

No Quadro 3, vejamos os verbetes para a unidade lexical *abobrinha*, sendo que os dois primeiros dicionários são os escolares e os seguintes, os gerais:

Quadro 3 – Verbetes de ‘abobrinha’ adaptados pelos autores

Dicionário	Definição
CA	a.bo.bri.nha <i>sf.</i> 1 A <u>abobrinha</u> é um fruto comestível da família da abóbora, de menor tamanho. 2 gíria quando falamos de coisas não muito sérias, ou dizemos bobagens, estamos falando ou dizendo, <u>abobrinhas</u> .
DAI	a.bo.bri.nha <i>substantivo feminino</i> 1. Legume comprido, verde-claro, que se come cozido. 2. <i>Gíria</i> Conversa sem sentido ou sem importância: <i>Em vez de estudar, passaram a tarde falando <u>abobrinhas</u>.</i>

Dicionário	Definição
AU	substantivo feminino. Bras. 1. Bot. Fruto verde da aboboreira, tão apreciado quanto o maduro na alimentação do homem. 2. Gír. Obsol. Cédula de mil cruzeiros antigos. 3. Gír. Bobagem, tolice: É pessoa inconsequente; só fala <i>abobrinha</i> .
HO	substantivo feminino 3 Regionalismo: Brasil. Uso: informal. conversa superficial 4 Regionalismo: Brasil. Uso: informal. afirmação que contém erro ou é, em si mesma, tola ou absurda 5 Regionalismo: Brasil. Uso: informal. Diacronismo: obsoleto. cédula de mil cruzeiros, de cor abóbora, com a efígie de Pedro Álvares Cabral, que circulou no Brasil, entre as décadas de 1940 e 1960
HOvir	substantivo feminino 3 <i>infrm.</i> conversa superficial 4 <i>infrm.</i> afirmação que contém erro ou é, em si mesma, tola ou absurda 5 <i>infrm.; obsl.</i> cédula de mil cruzeiros, de cor abóbora, com a efígie de Pedro Álvares Cabral, que circulou no Brasil, entre as décadas de 1940 e 1960.
CAvir	(a.bo.bri.nha) sf. 1. Variedade de abóbora, verde, us. na culinária. 2. Bras. Pop. Dito desprovido de conteúdo, sem importância (falar <i>abobrinha</i>); BOBAGEM; BESTEIRA [F.: <i>abobra</i> + <i>-inha</i> .]

Fonte: elaborado pelos autores a partir dos dicionários do *corpus*

A efemeridade da gíria faz com que seu uso esteja circunscrito a uma determinada época cujas condições de produzir os efeitos dela esperados se encontram na vontade dos falantes. Desse modo, quando os indivíduos deixam de usá-la, ela passa a ser considerada como uma lexia arcaica, antiquada, que caiu em desuso. À vista disso, encontramos em AU, HO e HOvir o sentido “cédula de mil cruzeiros antigos, de cor abóbora” para a entrada ‘abobrinha’ marcado como obsoleto, revelando ser um significado que os usuários da língua não mais o utilizam, sendo ultrapassado. Com efeito, os dicionários escolares CA e DAI não reproduzem esse sentido, o que demonstra acuidade dessas obras em inserir sentidos que sejam atuais e frequentes na língua.

Além disso, também encontramos um sentido marcado por ‘gíria’ que é contemporâneo e se refere a bobagens, tolices, algo sem importância, de pouca relevância, ditas por uma pessoa. Essa acepção consta nos dicionários escolares (CA e DAI), mas também é dicionarizado nos gerais AU, HO, HOvir e CAVir, que nos servem de parâmetros uma vez que são reconhecidos como modelos de referência e autoridade linguística para os brasileiros. No entanto, os gerais dicionarizam e marcam esse sentido de ‘abobrinha’ diferentemente: AU o etiqueta como ‘gíria’, seguindo os escolares; HO e HOvir rotulam-no como ‘informal’ e CAVir como ‘popular’. Ao se deparar com diferentes marcações, o consultante pode não compreender, de fato, o significado almejado e, assim, a obra dicionarística acaba por não atingir seu fim.

Essa evidente não padronização das marcas de uso corresponde às escolhas feitas por seus lexicógrafos a partir de sua visão de mundo. Logo, poderão considerar as variações linguísticas, e sua necessidade de marcação, conforme seu entendimento sobre as convenções sociais, regionais, estilísticas e históricas. Ao ser dicionarizada, a gíria ganha uma inserção na comunidade falante que a descaracteriza a ponto de, a nosso ver, precisar ser marcada por ‘informal’, ‘popular’ ou mesmo ‘coloquial’ em razão de sua ampliação. De fato, ‘abobrinha’ deixa de ser uma gíria utilizada apenas por jovens e passa a ser usada por vários outros grupos sociais.

Vejamos os verbetes relacionados a ‘bico’:

Quadro 4 – Verbetes de ‘bico’

Dicionário	Definição
CA	<u>bi.co</u> sm. 4 gíria Um <u>bico</u> é também um trabalho que uma pessoa faz de vez em quando, quando aparece: <i>Desempregado, ele vive de bicos.</i>
DAI	bi.co substantivo masculino 3. Gíria Trabalho provisório que se faz para ganhar dinheiro extra.

Dicionário	Definição
AU	substantivo masculino. 13. Bras. Fam. Pequenos ganhos avulsos e/ou tarefa ocasional que os possibilita; biscate, gancho, galho, viração. 14. Bras. Fam. Emprego subsidiário, pouco rendoso; gancho. 15. Bras. Gír. V. <i>cachaça</i> . (1). Gír. 1. Falar, dizer. 2. Delatar, denunciar; despejar o saco. 3. Confessar. 4. Dar (o atleta) mostras de cansaço. Baixar o bico. Bras. Comer ou beber em excesso. De bico. Gír. Fut. Com a ponta da chuteira. Molhar o bico. Bras. Pop. V. <i>embriagar</i> .
HO	substantivo masculino 11 Uso: informal. biscate.
HOvir	substantivo masculino 20 <i>infrm. m.q.</i> biscate (no sentido de 'serviço', 'ocupação', 'trabalho', 'remuneração')
CAvir	(bi.co) sm 14. Pop. Trabalho extra ou eventual; BISCATE: Trabalhava no banco, mas arranhou um <u>bico</u> durante a noite 16. Bras. Gír. <i>Cachaça</i> .

Fonte: Elaborado pelos autores a partir dos dicionários do *corpus*

A respeito de 'bico', CA e DAI marcam um de seus sentidos como uma forma de trabalho temporária para se obter dinheiro assim como em AU. No entanto, nas diferentes versões do Houaiss, verificamos novamente o uso da marca 'informal' – e nota-se uma tendência para a utilização desse tipo de marcação –, embora seu consulente precise entender que 'biscate' se refira a esse tipo de labor [somente na versão *on-line* há essa menção explícita – *infrm. m.q.* biscate (no sentido de 'serviço', 'ocupação', 'trabalho', 'remuneração') –, o que não ocorre na versão impressa – *infrm. m.q.* biscate]. Por fim, destaca-se que em CAvir esse sentido é ausente.

Consideremos 'biruta' (Quadro 5):

Quadro 5 – Verbetes de ‘biruta’

Dicionário	Definição
CA	bi.ru.ta a2g. 1 gíria Dizemos que alguém é <u>biruta</u> quando tem um comportamento irrequieto, amalucado: <i>Esse cara é meio biruta, gosta de andar de costas.</i>
DAI	bi.ru.ta substantivo feminino 2. Gíria Que é um tanto maluco. adjetivo de dois gêneros 3. Gíria Pessoa um tanto maluca.
AU	substantivo feminino substantivo de dois gêneros. 2. Bras. Gír. Pessoa irrequieten, amalucada. Adjetivo de dois gêneros. 3. Bras. Gír. V. <i>amalucado</i> (1).
HO	adjetivo de dois gêneros e substantivo de dois gêneros 3 Derivação: por analogia. Regionalismo: Brasil. Uso: informal. relativo a ou indivíduo sem norte, sem opinião ou vontade firme, ou que se deixa influenciar 4 Derivação: por analogia. Regionalismo: Brasil. Uso: informal. que ou quem é inquieto, de comportamento variável; amalucado.
HOvir	adjetivo e substantivo de dois gêneros (1979) 3 <i>p.ana.; B; infrm.</i> relativo a ou indivíduo sem norte, sem opinião ou vontade firme, ou que se deixa influenciar 4 <i>p.ana.; B; infrm.</i> que ou quem é inquieto, de comportamento variável; amalucado sinônimos
CAvir	(bi.ru.ta) sf. a2g. 2. Bras. Pop. Que é um tanto amalucado, doido. s2g. 3. Bras. Pop. Pessoa biruta (2). [F.: De or. obsc.]

Fonte: elaborado pelos autores a partir dos dicionários do *corpus*

Assim como verificado em ‘bico’, as marcas de uso utilizadas para ‘biruta’ nos vários dicionários comprovam que, enquanto CA, DAI e AU etiquetam os sentidos como ‘gíria’, os demais, HO e HOvir, os marcam como ‘informal’ e CAvir como ‘popular’. Depreende-se disso que, para os dicionários infantis (e também para AU), a marca ‘gíria’ deve remeter aos sentidos de itens lexicais criados por grupos sociais que os utilizam frequentemente. Já a equipe lexicográfica das versões de Houaiss considera esse uso como linguagem informal que, a princípio, restringia-se a grupos marginais, mas que, hoje, pode ser considerada estendida a outros grupos sociais e, portanto, de-

cidiu etiquetá-lo como informal. Por sua vez, em CAvir, a Lexikon (responsável por esse projeto lexicográfico) utiliza não apenas ‘gíria’, quando entende que corresponda a uma restrição a grupos sociais e/ou profissionais, mas também tabu, quando se refere ao sagrado ou ao proibido; por outro lado, utiliza as marcas ‘popular’ ou ‘vulgar’ para destacar, respectivamente, um uso corriqueiro e um grosseiro (por exemplo, acepção 3 de ‘pau’). Destaca-se ainda que ‘biruta’ figura sob formas homônimas em AU (‘biruta’), HO e HOvir (‘biruta’), mas o mesmo não se verifica nas obras escolares, fato esse omitido aos consultantes, mas que poderia contribuir para uma melhor reflexão sobre a língua portuguesa, máxime se possui sentidos marcados. Examine-mos, a seguir, ‘duro’ (Quadro 6):

Quadro 6 – Verbetes de ‘duro’

Dicionário	Definição
CA	<u>du.ro</u> a. 4 gíria Quando uma pessoa está sem dinheiro diz-se que está <u>dura</u> .
DAI	du.ro adjetivo 4. <i>Gíria</i> Sem dinheiro: <i>Está duro, não pôde comprar a passagem.</i>
AU	-----
HO	adjetivo 10 Derivação: sentido figurado. Regionalismo: Brasil. Uso: informal. sem dinheiro 11 Regionalismo: Brasil. Uso: informal. que está completamente ocupado Ex.: <i>praça d. de gente</i> 12 Rubrica: música. que desagrada ao ouvido Ex.: <i>acorde d.</i>
HOvir	adjetivo 12 <i>fig.; B; infrm.</i> sem dinheiro < <i>queria comprar um presente para a mãe, mas estava d.</i> > 13 <i>B; infrm.</i> que ainda é forte; resistente, vigoroso < <i>é d., apesar de idosa</i> > 14 <i>B</i> que sabe brigar; valente, agressivo 15 <i>B; infrm.</i> que está completamente (ou quase) ocupado < <i>praça d. de gente</i> >
CAvir	duro ¹ (<u>du.ro</u>) a. 9. Bras. Gír. Que está sem dinheiro, desprovido de recursos financeiros: <i>Quería ir ao cinema, mas estava <u>duro</u>.</i>

Fonte: elaborado pelos autores a partir dos dicionários do *corpus*

Curiosamente, AU é o único dicionário de nosso *corpus* a não registrar o sentido de ‘duro’ (indivíduo sem dinheiro). Essa ausência não se estende às outras obras, pois tanto os escolares quanto os outros três o marcam seja como ‘gíria’ ou ‘informal’. Constatamos que é comum encontrar frases-exemplo nas versões *on-line*, o que pode configurar em tendência lexicográfica para as obras virtuais. Assim como ocorrido em ‘biruta’, temos formas homônimas marcadas em AU (‘duro²’), HOvir (‘iduro’) e CAvir (‘duro¹’). Ao contrastarmos as obras, confirmamos a continuidade no uso gírico de ‘duro’ pelos dicionários virtuais, sendo que CAvir destaca ainda a seus consulentes que possui “mais de 818 mil verbetes, definições e locuções em permanente atualização. Um dicionário de crescimento infinito, sempre em interação com a língua portuguesa” (2021, *on-line*).

Analisemos ‘irado’ (Quadro 7):

Quadro 7 – Verbetes de ‘irado’

Dicionário	Definição
CA	i.ra.do a. 1 Alguém fica <u>irado</u> quando algo o deixa com muita raiva: <i>Estava irada por terem quebrado sua boneca. 2 gíria</i> Algo ou alguém é <u>irado</u> se é bacana, legal, bom ou interessante: <i>ganhou um presente irado do pai.</i>
DAI	i.ra.do adjetivo 1. Cheio de ira, de raiva: <i>O rapaz ficou irado com a notícia de que não teria aumento. 2. Gíria</i> Muito bom, legal: <i>Para os amigos, seu corte de cabelo ficou irado.</i>
AU	adjetivo. 1. Enraivecido, colérico, assanhado. 2. Bras. Gír. Legal, bacana: <i>Fui a uma festa irada.</i>
HO	-----
HOvir	-----
CAvir	(i.ra.do) a. 2. Fig. Extremamente agitado, revoltado (mar <u>irado</u>) 3. Gír. Muito bom, ou muito legal, ou muito bonito, ou muito interessante, ou muito excitante, ou muito moderno etc. (tênis irado; viagem irada; menino irado; PC irado); SINISTRO: <i>Ganhou um presente irado do pai.</i> [F.: Part. de <i>irar</i> .]

Fonte: elaborado pelos autores a partir dos dicionários do *corpus*

Em ‘irado’, um de seus sentidos (situação muito agradável, divertida) se encontra marcado por ‘gíria’ nos dicionários do *corpus*, à exceção das versões de Houaiss, as quais nem mesmo lematizam essa unidade lexical. Nas versões de Houaiss, o consulente é advertido sobre certas convenções presentes, dentre as quais modos e formas nominais podem receber tratamento diferenciado a depender de vários motivos (2009, p. XXXI). Assim, parece-nos que, caso fosse considerado um adjetivo e não um particípio, ‘irado’ poderia ter sido dicionarizado, quer como polissêmico, quer como homônimo.

Averiguemos ‘morar’ e ‘pronto’ (Quadros 8 e 9):

Quadro 8 – Verbetes de ‘morar’

Dicionário	Definição
CA	mo.rar vb. 4 gíria Morar é perceber, compreender. int.: <i>De metrô você vai mais rápido, morou?</i> [Conjug. quadro 1: morar.]
DAI	mo.rar verbo 3. Gíria Frequentar assiduamente um lugar: <i>Não o procure em casa: <u>mora</u> na praia.</i>
AU	verbo transitivo circunstancial. 3. Bras. Gír. Frequentar assiduamente um lugar: <i>Ela <u>mora</u> na casa da amiga, não sai de lá.</i> verbo transitivo indireto. 4. Residir, viver: “ <u>Morava</u> com um tio, chefe de esquadra reformado.” (Machado de Assis, <i>Várias Histórias</i> , p. 82.) 5. Bras. Gír. Entender, compreender; manjar; sacar: <i>Você <u>morou</u> no que eu disse?</i> verbo intransitivo. 6. Morar (4): <u>Mora</u> só. 7. Bras. Gír. Morar (5): <i>Não seja insistente, ela não quer nada com você, <u>morou</u>?</i> [Cf. <i>mourar</i> e <i>murar</i> .]
HO	verbo transitivo indireto 5 Regionalismo: Brasil. Uso: informal. frequentar (local) com assiduidade; visitar (local) com frequência Ex.: <i>ávido por leituras, mora na Biblioteca Nacional</i> transitivo indireto 6 Regionalismo: Brasil. Uso: informal. entender, compreender Ex.: <i>não morou naquilo que o chefe disse.</i>

Dicionário	Definição
HOvir	verbo 5 <i>t.i. (prep.: em); B; infm.</i> frequentar (local) com assiduidade; visitar (local) com frequência <ávido por leituras, mora na Biblioteca Nacional>
CAvir	(mo.rar) v. 3. Bras. Gír. Perceber, compreender; MANJAR; SACAR [tr. + em : <i>Morou no</i> que ele falou?: “ <i>Mora na</i> filosofia/ prá quê rimar amor e dor?...” (Monsueto e Arnaldo Passos , <i>Mora na filosofia</i>) 4. Bras. Fig. Visitar lugar com frequência [ta. : <i>Ele mora no bar da esquina</i>] [F.: Do lat. <i>morare</i> , por <i>morari</i> . Hom./ Par.: <i>mourar</i> e <i>murar</i> (vários tempos do v.).]

Fonte: elaborado pelos autores a partir dos dicionários do *corpus*.

Quadro 9 – Verbetes de ‘pronto’

Dicionário	Definição
CA	pron.to a. sm. 4 gíria Pronto é quem não tem dinheiro: Esse cara é um <i>pronto</i> .
DAI	pron.to adjetivo 4. Gíria Sem dinheiro, limpo.
AU	adjetivo 10. Bras. Gír. Diz-se do indivíduo sem dinheiro, ou pobre; areado, duro, estourado, fino, frito, limpo, liso, miqueado, quebrado, teso. substantivo masculino. 12. Bras. Gír. Indivíduo pronto (10); duro, frito, limpo, liso, teso.
HO	adjetivo e substantivo masculino Regionalismo: Brasil. Uso: informal. 8 que ou o que é pobre ou está sem dinheiro (diz-se de indivíduo) v advérbio 9 logo, em breve <i>Ex.: p. ele chegará</i>
HOvir	adjetivo e substantivo masculino (1911) <i>B; infm.</i> 9 <i>B; infm.</i> que ou o que é pobre ou está sem dinheiro (diz-se de indivíduo) <é um p.> <não vou comprar nada hoje porque estou p.>
CAvir	(pron. to) a.6. Bras. Gír. Que está sem dinheiro, que está quebrado, duro sm. 7. Bras. Gír. Indivíduo desprovido de dinheiro

Fonte: elaborado pelos autores a partir dos dicionários do *corpus*

A etiquetagem em ‘morar’ nos leva a questionar se, de fato, a marca ‘gíria’ é hodierna, e se, por conseguinte, está bem empregada. Parece-nos que não, pois três dos dicionários – AU, CA e CAVir – marcam uma de suas acepções, respectivamente, acepções 5, 4 e 3, para destacar o sentido de compreender algo ou determinada situação, o que nos remete a um modo de falar restrito à época da Jovem Guarda em que um de seus expoentes, Roberto Carlos, utilizava-a e que, hoje, já não é mais frequente, tendo deixado, portanto, de ser adotada. A própria frase-exemplo em CAVir corresponde a compositores das décadas de 1960-70. Já o outro sentido marcado (frequentar assiduamente) mostra-se contemporâneo.

Da mesma maneira que ocorre em ‘morar’, a etiquetagem de ‘gíria’ para ‘pronto’ parece-nos abranger outra época. Notamos, por exemplo, que há uma datação em HOvir que pode sugerir seu uso no início do século passado. Embora esse dicionário entenda e trate sentidos gíricos como informais, este é um caso que poderia servir para atestar o surgimento e estabelecimento desse significado, tendo em vista que não é o escopo desta pesquisa tratar da etimologia, mas focar na etiquetagem atual, restringindo-nos à ‘gíria’.

Essa perspectiva sobre aquilo que já não é mais dito, que deixou de ser escrito, quer dizer, a perda de sentido atualizado para um grupo social específico trouxe à baila, em 2021, um novo conflito linguístico de gerações, no caso, entre *millennials* e a geração Z. Essa questão torna-se interessante, neste ponto da discussão, para podermos demonstrar sua efemeridade: a geração Z (nascidos entre 1995 e 2010) passaram a usar a gíria *cringe*, verbo inglês, que se tornou adjetivo gírico neste ano para se referirem ao comportamento dos *millennials* (nascidos entre 1980 e 1994) como sendo antigo e ultrapassado, mas que, quase instantaneamente, caiu em desuso: “usar *cringe* é *cringe*”. À vista disso, observamos que realmente a gíria é um fenômeno par-

ticular a uma determinada época e que, fora dela, perde seu sentido original, justamente por estar descontextualizada.

Consideremos 'pau' (Quadro 10):

Quadro 10 – Verbetes de 'pau'

Dicionário	Definição
CA	sm. 1 Um <u>pau</u> é qualquer pedaço de madeira ou qualquer pedaço de árvore sem folhas que podemos segurar com uma só mão. 2 também chamamos de <u>pau</u> de uma bandeira o mastro dela. 3 São Paulo gíria Real (dinheiro): <i>Pode me emprestar 10 <u>paus</u>?</i>
DAI	pau substantivo masculino 1. Pedaço de madeira. 2. Qualquer madeira: <i>Fez uma casinha de <u>pau</u> para o cachorro.</i> 3. Gíria Reprovação em exame: Não estudou para prova, e levou o <u>pau</u> . [Plural: <i>paus</i>].
AU	substantivo masculino. 11. Bras. Gír. Reprovação em exame; bomba: <i>Levou <u>pau</u> em matemática.</i> 12. Afric. Bras. N.E. Pop. Qualquer árvore. 13. Bras. Jorn. O ponto de exclamação, entre os revisores. 14. Bras. Mar. G. Gír. Serviço de quarto ou de dia: <i>estar de <u>pau</u> das oito ao meio-dia.</i> 15. Pop. Real, ou escudo, ou outra moeda qualquer: <i>Paguei 20 <u>paus</u> por este livro.</i> 16. Bras. Chulo O pênis. 17. Bras. Gír. V. rolo ¹ (16). adjetivo de dois gêneros. Chutar o pau da barraca. Bras. Gír. V. <i>descer o morro</i> : “Irritado porque Pelé foi contra a sua sugestão de escalar Edmundo, em vez de Müller, para a seleção brasileira, Romário <u>chutou o pau da barraca</u> . ‘Pelé não tem que dar palpite.’” (Revista <i>Veja</i> , 16.03.1994.). Nem a pau. Bras. Gír. De modo nenhum; em absoluto; nem morto; nem amarrado, nem a tiro. Bras. Gír. 1. Ser sorteado para o exército. 2. Ir lutar em uma guerra.
HO	substantivo masculino 7 Regionalismo: Brasil. Uso: informal. reprovação em exame <i>Ex.: levar p. no vestibular</i> 8 Uso: informal ou tabuísmo. pênis.
HOvir	substantivo masculino 11 (1911) B; <i>infrm.</i> reprovação em exame <i><levar p. no vestibular> <ir ao p. na prova de matemática> <tomar p. no concurso></i> .

Dicionário	Definição
CAvir	sm. 3. Vulg. O pênis; CACETE 5. SE Gír. Design. genérica de uma unidade monetária: <i>Custou 200 paus</i> 6. Pop. Reprovação em exame: <i>Tirou 2,0 em matemática e levou pau.</i> Chutar o pau da barraca 1 Bras. Gír. Agir com irritação e grosseria, ger. como reação a uma situação; engrossar: <i>Ao ser criticado não se conteve, chutou o pau da barraca</i> e partiu para a briga. Nem a pau 1 Bras. Gír. De jeito algum: <i>Com essa chuva não saio nem a pau.</i> Pegar no pau furado 1 Bras. Gír. Ser convocado ou sorteado para o serviço militar. 2 Fazer o serviço militar, esp. para lutar numa guerra.

Fonte: elaborado pelos autores a partir dos dicionários do *corpus*

Verificamos que os sentidos de ‘pau’ são aqueles que mais são ressaltados pela marca ‘gíria’. A produtividade de significados dessa lexia ocorre em todos os dicionários de nosso *corpus*, sendo que o sentido de reprovação escolar está presente em todos, salvo em CA. Estranhamente, esse dicionário escolar não registra esse que é um sentido esperado para circular, sobretudo em um ambiente educacional.

Analisemos ‘zoar’ (Quadro 11).

Quadro 11 – Verbetes de *zoar*

Dicionário	Definição
CA	zo.ar int. 3 gíria E ainda é se divertir para valer, sem se preocupar com nada. int. <i>Nesse fim de semana, só quero zoar.</i> 4 gíria Zoar alguém ou com alguém é falar algo engraçado dessa pessoa para implicar, por brincadeira, com ela. td.: <i>Ele zoava os colegas para chamar atenção. Não gosto que fiquem me zoando.</i> ti.: <i>Não é verdade o que eu disse, estava apenas zoando com você.</i> [Conjug.: quadro 1: zoar. pres.: zôo, zoas, zoa etc.]
DAI	zo.ar verbo 3. Gíria Caçoar de alguém ou de alguma coisa; gozar: <i>Zoou (com) o amigo na frente da namorada.</i>

Dicionário	Definição
AU	verbo intransitivo. 5. Gír. Fazer troça, brincadeiras: <i>Depois da confusão, disse que estava apenas zoando.</i> Verbo transitivo direto. 6. Gír. Caçoar, zombar, de: <i>Zoou o primo quando o time dele levou uma goleada.</i> Verbo transitivo indireto. 7. Gír. Zoar (6): <i>Não deixou os meninos zoarem com o velho.</i> [Norm. é defect., unipess. nas acepçs. de 1 a 3. Pres. subj.: zoe (ô), zoes (ô), etc. Cf. Zóis, antr.]
HO	verbo transitivo direto, transitivo indireto e intransitivo 4 Regionalismo: Brasil. Uso: informal. fazer troça de; rir de alguém ou fazer-lhe uma brincadeira, por divertimento; caçoar, gozar Exs.: <i>zoava a pobre irmã na presença de todos; não me leves a mal, estava apenas zoando contigo; você não está falando sério, está zoando, não é?</i> intransitivo.
HOvir	verbo 3 int. produzir ruído semelhante ao dos insetos; zumbir, sibilar 4 t.d.,t.i.int. (prep.: com, de); B; infrm. fazer troça de; rir de (alguém) ou fazer-lhe uma brincadeira, por divertimento; caçoar, gozar <zoava a pobre irmã na presença de todos> <não me leves a mal, estava apenas zoando contigo> <você não está falando sério, está zoando, não é?> 5 int.; B; infrm. ocupar-se de maneira prazerosa; ir a algum lugar onde há divertimento; divertir-se <em vez de estudar, saiu para z> 6 int.; B infrm. promover confusão, desordem <aqueles baderneiros foram à discoteca para z> sinônimos ver sinonímia de <i>retumbar</i> .
CAvir	(zo.ar) v. 4. Bras. Gír. Divertir-se sem compromisso com nada [int.: <i>Sáiram à tarde só para zoar</i>] 5. Bras. Gír. Fazer pilhéria, caçoar [td.: <i>A garotada zoava o pobre mendigo: Você está a fim de me zoar, hem?</i>] [F.: Voc. onom.]

Fonte: elaborado pelos autores a partir dos dicionários do *corpus*

Finalizamos nossas análises com os sentidos etiquetados de ‘zoar’, sendo que três dicionários (CA, HOvir e CAvir) ressaltam o sentido de divertir-se sozinho sem que haja nenhum tipo de preocupação, mas todos os seis marcam o sentido negativo de caçoar alguém por divertimento, quer seja por ‘gíria’ ou ‘informal’, este último presente em muitas discussões em ambiente escolar quando se trata

de *bullying*, sendo, portanto, necessário fazer parte dos expedientes lexicográficos.

CONSIDERAÇÕES FINAIS

Nesta pesquisa, realizamos uma análise por meio do recorte de nove unidades lexicográficas que possuem sentidos dicionarizados e etiquetados pela marca de uso ‘gíria’. A partir disso, constatamos seu uso como estratégia para identificação e interação social, além de considerar que a variedade linguística não deve ser tomada como certa ou errada e que os dicionários acertam quando destacam esse uso por essa marca, pois é um registro histórico, uma “foto” da sociedade naquele momento em que o item lexical adquire o valor gírio.

Por ser um fenômeno sociolinguístico, concluímos que as lexias gírias se apresentam por duas perspectivas: (i) gíria grupal; (ii) gíria comum. O primeiro caso sugere uma restrição de uso, positiva ou negativamente, a um determinado grupo, neste recorte temos ‘biruta’, ‘duro’, ‘irado’, ‘morar’ e ‘pronto’, pois delimitam-nas no espaço-temporal, como analisado anteriormente. Por outro lado, as unidades lexicais consideradas que pertencem à perspectiva de gíria comum ocorrem devido a seu uso vulgarizado, quer dizer, que ultrapassou os limites do grupo que a utilizava, passando a ocupar os outros contextos sociais, são os nossos exemplos ‘abobrinha’, ‘bico’, ‘pau’ e ‘zoar’.

A partir da análise contrastiva entre os dicionários selecionados, evidenciamos que, por vezes, os itens lexicais gírios são destacados pelas marcas de uso ‘informal’ ou ‘coloquial’, recurso que, na verdade, entendemos que deveria ser, assim, padronizado, visto que, ao ser dicionarizada, a gíria perde seu caráter de variante particular a um número restrito de pessoas, com a finalidade de ser secreto, isto

é, oculto e cifrado. Essa barreira linguística é ultrapassada assim que a gíria passa a constar nas obras dicionarísticas, pois é a inconstância e a transitoriedade, que lhe são peculiares, que contribuem para a proposital incompreensão do público geral em benefício de um grupo em particular.

Em vista disso, para entender o fenômeno da gíria, deve-se levar em consideração que se trata de uma fotografia de um específico momento sócio-histórico em que determinado grupo fez uso de seu arcabouço lexical para se apartar, linguisticamente, da sociedade. Por isso a marca ‘informal’ pode substituir ‘gíria’, pois consegue recuperar o sentido de uma lexia que se apresente destituído de formalidade ou convencionalidade, da mesma maneira a marca ‘coloquial’ pode ser empregada no lugar de ‘gíria’ nos contextos em que os significados se aproximem da linguagem cotidiana.

Por fim, no verbete, esperam-se informações que façam referência, por exemplo, ao registro (no nosso caso, ‘gíria’) em que determinada palavra ou sentido sejam utilizados. Do ponto de vista geográfico, é difícil atestar que uma lexia corresponde a uma gíria, a não ser que seja realizada uma ampla investigação. Embora se possa afirmar que, predominantemente, gírias fazem parte da linguagem urbana, verificamos que a etiqueta ‘gíria’ é frequentemente usada pelos dicionários escolares, entretanto, como vimos, aspectos geográficos podem confirmar que os itens lexicais etiquetados por essa marca de uso podem não ser nacionalmente (re)conhecidos.

REFERÊNCIAS

AULETE, Caldas. **Aulete Digital**. Dicionário contemporâneo da língua portuguesa: Dicionário Caldas Aulete, online. Rio de Janeiro: Lexikon Editora digital, 2021. Disponível em: <http://www.aulete.com.br/>. Acesso em: 23 set. 2021.

BAGNO, Marcos. **Preconceito Linguístico**: o que é, como se faz. 49. ed. São Paulo: Editora Loyola, 2007.

BIDERMAN, Maria Tereza Camargo. **Dicionário ilustrado de português**. 2. ed. São Paulo: Ática, 2009.

BIDERMAN, Maria Tereza Camargo. Glossário. **Alfa**, São Paulo, v. 28 (supl.), p. 135-144, 1984.

BRAGA, Rita de Cássia Espechit; MAGALHÃES, Márcia A. Fernandes. **Fala Brasil!** Dicionário ilustrado da língua portuguesa. Belo Horizonte: Dimensão, 2011.

BRASIL. Ministério da Educação. Secretaria de Educação Básica. **Com direito à palavra**: dicionários em sala de aula (PNLD 2012: Dicionário). Brasília: Ministério da Educação, Secretaria de Educação Básica, 2012.

COELHO, Izete Lehmkuhl; GÖRSKI, Edair Maria; SOUZA, Christiane Maria Nunes de; MAY, Guilherme Henrique. **Sociolinguística**. Florianópolis: LLV/CCE/UFSC, 2010. Disponível em: https://ppglin.posgrad.ufsc.br/files/2013/04/Sociolingu%C3%ADstica_UFSC.pdf. Acesso em: 01 out. 2021.

COELHO, Izete Lehmkuhl; GÖRSKI, Edair Maria; SOUZA, Christiane Maria Nunes de; MAY, Guilherme Henrique. **O estudo da linguagem no contexto social**: para conhecer sociolinguística. São Paulo: Contexto, 2015.

FERREIRA, Aurélio Buarque de Holanda. **Dicionário Aurélio ilustrado**. Curitiba: Positivo, 2008.

FERREIRA, Aurélio Buarque de Holanda. **Novo Dicionário Eletrônico Aurélio versão 7.0**®. 5. ed. Curitiba: Editora Positivo Informática LTDA, 2010.

GEIGER, Paulo. (org.). **Caldas Aulete** – Dicionário escolar da língua portuguesa. 3. ed. São Paulo: Globo, 2011.

HOUAISS, Antônio. **Dicionário Eletrônico Houaiss da Língua Portuguesa versão 1.0**. Rio de Janeiro: Editora Objetiva, 2009.

HOUAISS, Antônio. **Dicionário Houaiss da Língua Portuguesa**. UOL. Rio de Janeiro: Editora Objetiva, 2021. Disponível em: https://houaiss.uol.com.br/corporativo/apps/uol_www/v5-4/html/index.php#0. Acesso em: 24 set. 2021.

KRIEGER, Maria da Graça. Políticas públicas e dicionários para escola: o Programa Nacional do Livro Didático e seu impacto sobre a Lexicografia Didática. **Cadernos de Tradução**, Florianópolis, v. 2, n. 18, p. 235-252, 2006.

NOVOA, Alberto Barahona. La marcación en la lexicografía actual. **Revista Káñina**, v. 29, p. 29-52, 2005.

PATRIOTA, Luciene Maria. **A gíria comum na interação em sala de aula**. São Paulo: Cortez, 1998.

PONTES, Antônio Luciano. **Dicionário para uso escolar**: o que é como se lê. Fortaleza: EdUECE, 2009.

PORTO DAPENA, José Álvaro. **Manual de técnica lexicográfica**. Madrid: Gredos, 2002.

PRETI, Dino. Dicionários de gíria. **Alfa**, São Paulo, v. 44, p. 57-73, 2000.

ULHÔA, Martha Tupinambá de. Ele canta lindamente as dores e os amores: gestualidade musical nas canções de Roberto Carlos. *In*: ULHÔA, Martha Tupinambá de; PEREIRA, Simone Luci (Org.). **Canção romântica**: intimidade, mediação e identidade na América Latina. Rio de Janeiro: Folio Digital - Letra e Imagem, 2016. p. 119-144.

VILARINHO, Michelle Machado de Oliveira. **Proposta de dicionário informatizado analógico de língua portuguesa**. 2013. 307 f. Tese (Doutorado em Linguística) – Universidade de Brasília, Brasília, 2013.

WELKER, Herbert Andreas. Marcas de uso. *In*: WELKER, Herbert Andreas. **Dicionários**: uma pequena introdução à lexicografia. 2. ed. revista e ampliada. Brasília: Thesaurus, 2004. p. 130-149.

WELKER, Herbert Andreas. **Panorama Geral da Lexicografia Pedagógica**. Brasília: Editora Thesaurus, 2008.

MEMÓRIAS DA “VIAGEM” DA CONSTRUÇÃO DO DICIONÁRIO HISTÓRICO DO PORTUGUÊS DO BRASIL (DHPB) E SUAS REPERCUSSÕES LEXICOGRÁFICAS

Clotilde de A. Azevedo Murakawa

INTRODUÇÃO



título deste texto reflete o pensamento de quem o escreve: a história da construção do *Dicionário Histórico do Português do Brasil – séculos XVI, XVII e XVIII* (DHPB) que pode ser “contada” como um relato de viagem, gênero narrativo tão representativo no período dos descobrimentos portugueses. Acrescente-se a esse pensamento o de Ramón Menéndez Pidal (1869-1968), filólogo e linguista espanhol, que assim que posiciona sobre a obra lexicográfica:

En fin, el Diccionario, al dar de cada palabra una caracterización precisa, una biografía esquemática que muestre las cambiantes o indecisas actitudes de la vida, representará el habla, no en reposo de autorizada estabilidad, sino en movimiento de avance; será como una fotografía instantánea del idioma en actitud dinámica, representando al vivo la dirección de su movimiento. No será un panteón consagrado al culto del bien decir, sino guía en una etapa del inacabable camino que va recorriendo el lenguaje.

Tal dicionário, fotografia instantânea, el mismo día que acaba de imprimir-se comienza a quedar anticuado, pero llevará en sí siempre, en las explicaciones del pasado y en la exacta descripción del presente, la razón de ser de las innovaciones futuras (Menéndez Pidal, 1961, p. 147).¹

A partir, portanto, da metáfora da fotografia instantânea de Menéndez Pidal, a forma como entendo o DHPB, passo a me valer do gênero relato de viagem, para resgatar o caminho percorrido pelo projeto desse dicionário, com suas turbulências e calmarias, da sua construção até a almejada publicação *online* da obra, em 2021, que transformou os 19 volumes originais impressos em tamanho A4 e encadernados com capa plástica em espiral, com seus 10.470 verbetes e 11.510 páginas, em uma obra de referência única, disponível na WEB a todos os interessados². Uma vez publicado, o DHPB torna-se uma “fotografia instantânea” do repertório lexical da língua portuguesa veiculado no período colonial brasileiro, como sugere o mapa das capitânicas hereditárias tão significativas no processo histórico de configuração desse período da história do Brasil.

¹ “Enfim, o Dicionário, ao dar uma caracterização precisa de cada palavra, uma biografia esquemática que mostre as mutáveis ou indecisas atitudes da vida, representará a fala, não no repouso de uma estabilidade autorizada, mas sim em um movimento de avanço; será como uma fotografia instantânea do idioma em atitude dinâmica, representando ao vivo a direção de seu movimento. Não será um panteão consagrado ao culto do bem dizer, mas um guia em uma etapa do inacabável caminho que a linguagem percorre. Tal dicionário, fotografia instantânea, no mesmo dia em que é publicado começa a ficar antiquado, mas levará sempre em si, nas explicações do passado e na exata descrição do presente, a razão de ser das inovações futuras” (Trad. Odair Luiz Nadim).

² O DHPB pode ser acessado pelo seguinte endereço: <https://dicionarios.fclar.unesp.br/dhpb/>.

Figura 1 – Mapa das capitânicas hereditárias – Brasil Colônia



Fonte: DHPB. Disponível em: <https://dicionarios.fclar.unesp.br/dhpb/>. Acesso em: 18 set. 2023.

Divido esta narrativa em cinco etapas, cronologicamente organizadas, para que se possa ter a exata dimensão e a importância do DHPB no cenário da história do português do Brasil e, por extensão, da lexicografia brasileira.

1. A INTENÇÃO DA VIAGEM...: a formatação da proposta do dicionário e a submissão ao Edital MCT/CNPq nº 01/2005/Institutos do Milênio

Em novembro de 2005, Maria Tereza Biderman me telefonou para falar sobre o Edital Institutos do Milênio do CNPq (Conselho Nacional de Pesquisa) e do seu interesse em submeter um projeto propondo a construção de um dicionário histórico do português do Brasil que reunisse um repertório lexical, a partir de uma base de dados formada por textos produzidos nos três séculos da colonização brasileira. Esse dicionário registraria o léxico representativo do início da língua portuguesa no Brasil.

Ainda durante a mesma ligação telefônica, ela me consultou se poderia colocar o meu nome como vice-coordenadora do projeto, convite que agradei e aceitei prontamente. Maria Tereza, minha orientadora de Mestrado, sabia do meu interesse pelos estudos lexicográficos do português em sua perspectiva histórica.

O projeto submetido ao MCT/CNPq nº 01/2005 sustentava-se em argumentos como a não existência de uma obra lexicográfica que descesse o vocabulário do Português Brasileiro antes do século XX, razão pela qual o projeto submetido pretendia criar e produzir um dicionário abrangendo o vocabulário dos primeiros séculos da história e da formação da sociedade brasileira. Para tanto, seriam coletados textos publicados e manuscritos brasileiros ou sobre o Brasil, representativos dos séculos XVI, XVII e XVIII. Seleccionadas e coletadas essas fontes de referência, seriam criados os recursos tecnológicos necessários para dar forma digital a manuscritos disponibilizados em vários suportes, incluindo os textos impressos. Assim, paulatinamente, seria criada uma base textual informatizada do material recolhido que

se configuraria como banco de dados da qual seriam extraídos acervos lexicais com dois tipos de dados: 1) os *index verborum* para subsidiar a elaboração da nomenclatura do dicionário e, 2) as concordâncias dos lemas selecionados cuja extração contaria com o auxílio de outras ferramentas computacionais. Assim, segundo o projeto submetido, a partir desse repertório de concordâncias e de outros dados oriundos de consultas bibliográficas, seriam redigidos os verbetes, documentando os significados e os usos registrados na base textual construída com o objetivo mais amplo de produzir uma descrição histórica do vocabulário do Português Brasileiro nos primeiros estágios de sua formação que seria registrada em um Dicionário Histórico do Português do Brasil (Processo nº 420040/2008-0/CNPq). Aprovado o projeto e recebida a verba liberada, Maria Tereza deu início aos trabalhos necessários para concretizar o fim proposto.

2. E A VIAGEM INICIAVA O SEU PERCURSO...: a formação da equipe e a constituição do *corpus*

No ano de 2006. Maria Tereza faz o convite para docentes das mais variadas faculdades e universidades do Brasil para integrarem o grupo de trabalho que deveria atuar no projeto. Não relaciono aqui os nomes de todos os que participaram e deram sua contribuição para o projeto para não correr o risco de me esquecer de alguém. A relação completa de todos os componentes do grupo está disponibilizada tanto na versão impressa, em seu 1º volume, quanto na versão *online*, na parte da *Apresentação*. Muitos pesquisadores deixaram o projeto ao longo dos anos por questões de ordem vária, mas todos estão mencionados na versão final do dicionário.

Desta forma, na segunda etapa do projeto, foram feitas as aquisições de equipamentos como computadores, scanners, impressoras, máquina fotográfica, como também o material de consumo necessário para a instalação do projeto no Laboratório de Lexicografia da FCLAr – Faculdade de Ciências de Letras, UNESP/Araraquara.

Composta a parte material necessária, teve início o levantamento dos textos para a construção do banco de dados. Enquanto os colaboradores encaminhavam para a coordenação do projeto textos em diversos formatos para comporem o banco de textos, Maria Tereza e eu iniciávamos a busca de textos em bibliotecas e arquivos como: Biblioteca Mario de Andrade (SP), Arquivo Público do Estado de São Paulo, Instituto de Estudos Brasileiros (USP), Biblioteca Nacional do Rio de Janeiro, Gabinete Português de Leitura do Rio de Janeiro e Salvador, Museu da Inconfidência de Ouro Preto, Instituto Histórico e Geográfico de São Paulo e Belo Horizonte. Em Portugal, a Biblioteca Pública de Évora foi pesquisada e de lá foram trazidos alguns textos.

Alguns docentes enviaram textos pertencentes ao Mosteiro de São Bento de Salvador, ao Arquivo Público e ao Museu Emilio Goeldi no Pará e ao Instituto Histórico e Geográfico de Mato Grosso, além da busca feita por meio da rede de bibliotecas da UNESP, a partir da biblioteca da Faculdade de Ciências e Letras de Araraquara.

Embora no título do DHPB estejam registrados os séculos XVI XVII e XVIII, a procura dos textos se deu também no 1º decênio do século XIX por sugestão do Professor de História da Universidade de São Paulo/USP e renomado pesquisador Istvam Iancsó que ponderou sobre a necessidade de se estender o período de coleta até a vinda da Família Real portuguesa para o Brasil, em 1808, data que delimitou o fim do período colonial.

Desta forma, ao longo do ano de 2006, foi dada continuidade à busca de textos, tarefa que culminou com a formação de um acervo de textos que resultou em 23.858 páginas escaneada, com um total de 7.492.472 ocorrências.

O banco de dados que subsidiou a produção do DHPB começou, assim, a ser construído a partir da “carta de achamento do Brasil”, de Pero Vaz de Caminha, escrita em 1º de maio de 1500, documento que deu início à base de textos. Como já pontuado, a data-limite estabelecida para a seleção dos textos a serem inseridos na base de dados foi o ano de 1808, quando da vinda da Família Real portuguesa para o Brasil. Foram selecionados documentos dos três séculos considerados (XVI, XVII e XVIII), reunindo um conjunto representativo de vários gêneros e natureza de textos, a saber: obras dos missionários viajantes, na sua maioria jesuítas que vieram para o Brasil em missão catequética e aqui se fixaram; diários de navegação, como o de Pero Lopes de Sousa, irmão de Martim Afonso de Sousa; cartas de sesmarias; roteiros descritivos da flora e fauna brasileiras; descrições geográficas; cartas e sermões do Pe. Vieira, pregados aqui no Brasil, e de outros oradores sacros que para aqui vieram e que tiveram sua correspondência reunida em obras esparsas; obras e documentos que tratam do Estado do Grão Pará, durante a era pombalina; cartas comerciais trocadas entre comerciantes da colônia com outros de Portugal; obras sobre a nobiliarquia paulistana; atos de câmaras municipais; anais de câmaras de diversos municípios brasileiros; documentos cartoriais; autos de devassas feitos durante a Inconfidência Mineira; processos; inventários; testamentos; alvarás; posturas; bandos; atos de doações de terras, casas e terrenos; cartas de ofício; patentes; cartas dos governadores gerais; provisões; documentos forenses; estatutos de sociedades; constituições dos bispados do Brasil; regimentos militares; obras

sobre medicina, farmácia, agricultura, mineração, além da produção literária do barroco e arcadismo.

Os textos, na sua quase totalidade, foram digitalizados e submetidos a processamento informático. Desta forma, para que o conjunto selecionado e anteriormente relacionado pudesse ser inserido ao banco de dados, houve a necessidade de submissão dos textos a um tratamento informático, com a finalidade de deixá-los em formato adequado para que pudessem ser trabalhados com o auxílio do programa computacional Philologic, para esse fim adaptado. A ordem obedecida nesse processo foi a seguinte:

- a) escaneamento dos textos e edição das imagens;
- b) organização das pastas – uma para cada obra;
- c) transferência das imagens de textos TIFF (Tagged Image File Format) para Doc, a partir da leitura ótica (Optical Character Recognition – OCR) e devida correção pelo programa ABBYY Fine Reader;
- d) inclusão da ficha catalográfica nos textos corrigidos;
- e) conversão do material escaneado para arquivos texto (TXT);
- f) marcação XML (eXtensible Markup Language);
- g) inserção dos textos no programa Philologic.

Depois de escaneados e revisados os textos foram enviados ao Instituto de Ciências Matemáticas e da Computação (ICMC) da USP de São Carlos, para a sua devida inserção no banco de dados, elegendo-se o programa computacional Philologic como suporte para a extração das palavras do banco de textos.

A documentação escaneada teve sua imagem arquivada na Seção Técnica de Informática da FCLAr/UNESP/Araraquara.

Cumprir registrar que o trabalho relatado nesta etapa do projeto transcorreu de dezembro de 2005 a junho de 2007. Ao longo desse período, Maria Tereza realizou uma reunião em Araraquara com todos os participantes do projeto, para discutir diversos aspectos da sua execução.

3. MOMENTOS DE TRISTEZA, ANGÚSTIA E INCERTEZAS...

A partir de 2007, o banco a que denominei banco I estava pronto para o início da produção dos verbetes. Esse banco, conforme mencionado anteriormente, contém 7.492.472 ocorrências e estava alocado no ICMC/USP com acesso restrito à coordenação do projeto e à equipe de redatores.

Esta fase do projeto foi a mais difícil de todas, pois em junho desse mesmo ano Maria Tereza foi submetida a uma cirurgia no cérebro que a impossibilitou de continuar em atividade. Entretanto, todo o trabalho no Laboratório de Lexicografia/FCLAr/UNESP continuou sendo executado com os alunos que integravam o grupo.

O projeto entrou assim numa fase de incertezas e sem perspectivas de continuidade. Enquanto Maria Tereza pôde, continuou efetuando o pagamento para os alunos que atuavam no laboratório e alguns deles receberam bolsa da Reitoria da UNESP, o que garantiu a continuidade do trabalho de digitalização, escaneamento, correção e revisão dos textos. Nesse período, os pesquisadores ainda não tinham recebido as listas de palavras para a redação dos verbetes.

Momentos de grande tristeza, angústia, insegurança tomaram conta de mim e dos membros do projeto, particularmente os que atuavam no laboratório. Mesmo assim o trabalho seguiu seu ritmo com a esperança de que tudo retornasse à normalidade. Infelizmente, Maria Tereza veio a falecer em maio de 2008.

As dificuldades foram se acentuando, pois a verba liberada pelo CNPq estava sob responsabilidade de Maria Tereza. Mesmo assim o projeto não parou. Pagamentos e auxílio alimentação para os estagiários, assim como material de consumo do laboratório, foram bancados pela direção da FCLAr/UNESP, durante a gestão do Prof. Claudio Benedito Gomide de Souza e do Prof. José Luís Bizelli, até o recebimento da posição do CNPq sobre a continuidade dos trabalhos previstos no cronograma do projeto.

Juntamente com o Prof. João Moraes Pinto Jr., responsável pela Informática no âmbito de Laboratório de Lexicografia, reuni todos os comprovantes de pagamentos por serviços prestados e compras feitos até a data de falecimento de Maria Tereza e enviamos ao CNPq como comprovação do trabalho executado até aquela data. A verba depositada no Banco do Brasil em nome de Maria Tereza retornou para o CNPq.

Ressalto, aqui, que tanto a ausência de perspectiva de continuação do projeto quanto a falta de verba não permitiram que o trabalho parasse e que houvesse dispersão da equipe em especial dos alunos colaboradores.

Somente em agosto de 2008, recebi uma ligação telefônica de uma funcionária do CNPq me informando que eu iria assumir a coordenação do projeto e que seria aberta nova conta em meu nome para viabilizar o recebimento da verba restante.

Esta foi, com certeza, a etapa mais sofrida vivenciada pelo projeto até a conclusão do dicionário.

4. CONTINUIDADE DO PERCURSO: a redação dos verbetes...

A partir do mês de setembro de 2008 os verbetes do DHPB começaram a ser redigidos. Listas de unidades lexicais passaram a ser extraídas do banco de textos com auxílio do programa UNITEX 2.0 e enviadas aos redatores. Esse foi um período bastante intenso e de muitas discussões sobre os procedimentos lexicográficos a serem adotados na construção dos verbetes, tanto no que se referia à definição, quanto às demais informações que deveriam constar na microestrutura dos verbetes.

Tendo em vista os inúmeros impasses que foram surgindo, houve a necessidade de se tomar por orientação as diversas teorias e metodologias propostas pela Lexicografia para a construção de um dicionário e, no caso específico, de um dicionário histórico.

A riqueza de dados oferecida pelo banco de textos levou a equipe a elaborar critérios específicos para a construção dos verbetes, levando em conta os diversos contextos que o banco oferecia para cada caso. Tais procedimentos teóricos e metodológicos integram a parte introdutória das versões impressa e *online* do DHPB.

Com conhecimento dos critérios para redação, os docentes redatores começaram a construir os verbetes a partir de cada conjunto de palavras recebidas. O pagamento dos redatores era realizado na proporção em que os lotes de verbetes eram recebidos pelo laboratório.

Estabeleceu-se que, à medida que os verbetes prontos eram recebidos pelo laboratório, as estagiárias Carolina Domladovac-Silva e Ive Toder faziam uma primeira revisão do material e, na sequência, a mim enviados para as devidas e necessárias correções. A verificação dos verbetes sempre teve como orientação o registrado no banco de textos, razão pela qual as revisões se configuraram como minuciosas e demoradas. Devolvidos ao laboratório, os verbetes prontos eram arquivados e armazenados em ordem alfabética. Esse processo de revisão e correção teve início a partir de agosto de 2008.

Paralelamente ao processo de redação e revisão, resolvi, por minha conta e risco, continuar ampliando o banco de dados, por ter percebido que o século XVII não tinha a mesma representatividade de textos que o século XVIII. Pouco havia do período de seiscentos em termos de textos e precisaria ser representado tanto quanto os demais. Foi construído, então, o banco II com um total de 2.049.249 ocorrências a partir de 8.500 páginas escaneadas.

Na organização do banco II, não houve mais necessidade de envio dos textos processados para o ICMC da USP, de São Carlos. Assim, todo o processamento e a construção do banco foram realizados no próprio Laboratório da Lexicografia da FCLAr/UNESP.

Importante registrar que, até esse momento, o projeto estava buscando o cumprimento do prazo estabelecido pelo CNPq para a sua conclusão: novembro de 2010. Entretanto, constatei que não seria possível a finalização dos 10.000 verbetes previstos por Maria Tereza no prazo estipulado no projeto original.

Desta forma, em reunião organizada pelos representantes do CNPq, em São Paulo, com a participação dos coordenadores dos diversos projetos contemplados pelo Edital Institutos do Milênio, tive oportunidade de me encontrar com um amigo de família, Prof. Dr.

José Galizia Tundisi, ex-presidente do CNPq. Falando com ele sobre os problemas ocorridos durante a execução do projeto, ele sugeriu que eu solicitasse ao CNPq um termo aditivo, com ampliação do prazo para o fechamento do projeto em 2012. Assim, seguindo essa orientação, encaminhei ao CNPq a solicitação que foi aprovada e o termo aditivo publicado no Diário Oficial da União de 11 de outubro de 2008 ampliou o prazo para conclusão do projeto para até 23 de novembro de 2012 e a prestação de contas para fevereiro de 2013.

O Laboratório de Lexicografia comemorou a prorrogação e os trabalhos tiveram continuidade em ritmo mais acelerado, pois em novembro de 2012 o Projeto DHPB deveria encerrar a sua movimentação financeira junto ao Banco do Brasil. E, em fevereiro de 2013, o projeto concluído deveria ser enviado ao CNPq.

Dentro do prazo estabelecido, o acordo firmado foi cumprido e os exemplares do DHPB enviados por correio ao CNPq. Como já mencionado anteriormente, 19 volumes impressos em A4 com 11.510 páginas contendo 10.470 verbetes seguiram como comprovante do projeto concluído, além da conclusão de um banco de textos de 9.541.721 ocorrências e 32.358 páginas escaneadas, devidamente armazenado no Laboratório de Lexicografia da FCLAr/UNESP.

Devo ressaltar, entretanto, que para a redação dos 10.470 verbetes que compõem o DHPB foi utilizado apenas o banco I. Logo, a exploração do banco II representa novos e futuros desafios.

Cumpridas as exigências legais, restava a prestação de contas. Durante os meses de fevereiro e março do ano subsequente foi reunida toda a documentação para comprovar mês a mês as despesas efetuadas. Saldadas as contas, devolvi ao CNPq o que tinha ficado da verba enviada para meu nome. Essa prestação de contas foi feita por Carolina Domladovac-Silva que tinha o conhecimento necessário

para efetivá-la no formato *online*. Em abril de 2013 as contas foram aprovadas sem ressalvas.

5. CHEGADA A BOM PORTO: desafios para a publicação do DHPB e suas repercussões...

Finalizado o projeto, procedeu-se ao levantamento de todos os livros adquiridos durante a execução das diferentes etapas da pesquisa, para fins de envio à Biblioteca da Faculdade de Ciências e Letras de Araraquara. Um volume representativo de dicionários não existentes na biblioteca passou a integrar o acervo. Isso sem contar as obras de diversos gêneros que foram adquiridas para serem escaneadas e digitalizadas para integrarem o banco de dados.

Inicia-se uma nova etapa a partir de abril de 2013 com o objetivo de publicar o DHPB. Minha primeira solicitação foi para o diretor da editora da UNESP que elogiou o trabalho, mas informou que não haveria condição de publicação. Apesar do empenho do Prof. Claudio César de Paiva, vice-diretor, e depois diretor da FCLAr, nada se conseguiu.

Muito embora não houvesse uma divulgação oficial e não estivesse permitida a consulta ao DHPB, eu recebia com muita frequência pedidos para consultas. Mas, como o dicionário não poderia ser liberado ao público por questões de segurança dos dados do banco de textos e do próprio DHPB, busquei atender a todos, fazendo as pesquisas solicitadas e enviando a cada um em particular³. E deste modo consegui ir divulgando, aos poucos, o nome e a importância do dicionário aos pesquisadores no Brasil e no exterior.

³ Em todos esses casos, o DHPB constou das referências bibliográficas dos trabalhos e, conseqüentemente, era divulgado para a comunidade acadêmica.

Depois de alguns anos sem se ter perspectivas de publicação do DHPB, independente do formato, surgiu a oportunidade de sua publicação *online*, graças ao empenho do Prof. Odair Luiz Nadin, do Departamento de Letras Modernas da FCLAr/UNESP e do Prof. José Machado Lopes Filho, especialista em Linguística de *Corpus* e com Mestrado na área, pela USP, que buscaram colocar na WEB o DHPB tal qual estava impresso em A4, quando do seu envio ao CNPq, publicação acompanhada do *corpus* que lhe deu suporte.

Devo, aqui, destacar o auxílio financeiro fornecido pela FCLAr/UNESP, durante a gestão do Prof. Dr. Jean Cristtus Portela, que também se empenhou para que o dicionário no formato *online*, fosse disponibilizado no site da Faculdade de Ciências e Letras de Araraquara: <http://dicionarios.fclar.unesp.br>. A Figura 2, na sequência espelha a página inicial do DHPB.

Figura 2 – Espelho da página inicial do Dicionário Histórico do Português do Brasil - DHPB



Fonte: DHPB: <https://dicionarios.fclar.unesp.br/dhpb/>. Acesso em: 18 set. 2023.

Concluo este relato de “viagem” retomando o pensamento inicial do texto de que a construção do *Dicionário Histórico do Português do Brasil – séculos XVI, XVII e XVIII* foi uma “viagem” iniciada em 2005 e concluída em setembro de 2021, quando essa obra de referência, inédita, foi oferecida à comunidade acadêmica interessada no estudo da língua portuguesa, em especial no seu repertório lexical que documenta a história do Brasil Colônia.

Uma obra dessa natureza foi deixando, ao longo do caminho de sua construção, contribuições lexicográficas importantes que poderão servir de parâmetro para pesquisas futuras. Relaciono, a seguir, alguns dos legados do DHPB para a Lexicografia Portuguesa:

1) *O DHPB é um dicionário documental.* Os verbetes foram redigidos a partir dos contextos extraídos do banco de textos com o devido registro da referência bibliográfica. Não há exemplos criados pelos redatores e as definições levaram em consideração os contextos evidenciados no banco de textos. É o caso, por exemplo, do lema *contagem* que em uma de suas acepções é definido como “alfândega entre as capitanias”, significado obtido por meio dos contextos em que a unidade léxica *contagem* figura no banco de dados (Figura 3).

Figura 3 – Verboete *contagem* em sua segunda acepção

2. Posto de fiscalização e arrecadação, especializado na cobrança de tributos incidentes sobre animais e outras mercadorias em trânsito de uma capitania para outra.

Da mesma sorte o criador de gado q.^a tem o escravo p.^o o beneficiar, e paga o dizimo delle, e alem deste nas **contages** paga huma outava por cada cabeça de gado, e por cavallo, a duas outavas por cabeça, e alem destes tributos paga o mesmo q.^o como paga o Mineiro por cada escravo. CAETANO LOPES LAVRE (1896) [1741], VI - IMPOSTOS NA CAPITANIA MINEIRA [A00_0779 p. 304].

Fonte: DHPB. Disponível em: <https://dicionarios.fclar.unesp.br/dhpb/>. Acesso em: 18 set. 2023.

2. *Teoria e metodologia lexicográficas empregadas de acordo com os objetivos do dicionário.* A tipologia da definição: gênero próximo e diferença específica; sinonímia; antonímia e descrição, chegando à definição enciclopédica, quando necessária, foram os modelos mais utilizados para a definição dos substantivos, adjetivos e verbos que compuseram a nomenclatura do dicionário;

3. *Informações constantes no verbete:* a partir do lema, vem a classe gramatical, os significados do lema, sempre a partir do contexto, ou seja, o que o banco de textos registra e a 1ª datação. Essas são informações obrigatórias que constam de todos os verbetes. Também figuram nos verbetes outras informações consideradas facultativas, pois dependem de sua presença ou não no banco de dados: variantes, sentido figurado, locuções, expressões sintagmáticas e informação enciclopédica. Para todas as informações obrigatórias e facultativas o contexto e a informação bibliográfica estão registrados no corpo do verbete;

4. *Registro da 1ª datação.* Essa datação se refere ao documento mais antigo do banco que registra a palavra-entrada. O documento mais antigo é a Carta de Pero Vaz de Caminha. Essa retrodatação tem fornecido informações importantes para a equipe do Instituto Antonio Houaiss para a revisão de número significativo de palavras do Dicionário da Língua Portuguesa que constavam com o registros de datas mais recentes.

Ao registrar o documento com data mais antiga, até mesmo fatos da história do Brasil podem ser revisados. Foi o que observei em uma notícia veiculada pelo jornal o Estado de São Paulo de 11 de ou-

tubro de 2021 que mencionava, em uma reportagem, que o *bairro da Saúde* no Rio de Janeiro é dos melhores lugares para se conhecer e que o bairro é povoado desde o século XVII, tendo recebido esse nome em 1742; menciona, ainda, que é nele que se localiza o *Cais do Valongo*, local de desembarque de pessoas escravizadas, inaugurado em 1811. Entretanto, a menção a esse cais, eu encontrei no banco de dados do DHPB em texto de um relatório do Marquês de Lavradio de 1799. Transcrevo, a seguir, a passagem do texto, onde consta essa data:

Foi a resolução ordenar que todos os escravos que viessem n'estas embarcações, logo que dessem sua entrada na Alfandega pela parte do mar, tornassem a sahir, e embarcassem para o sitio chamado **Vallongo**, que é no suburbio da cidade, separado de toda a comunicação; que alli se aproveitassem das muitas casas e armazens que alli ha para os terem; e que áquelles sitios fossem as pessoas que os quizessem comprar, e que os compradores nunca podessem entrar de quatro a cinco na cidade, MARQUEZ DO LAVRADIO (1863) [1799], *RELATORIO DO MARQUEZ DE LAVRADIO, VICE-REI DO RIO DE JANEIRO, ENTREGANDO O GOVERNO A LUIZ DE VASCONCELLOS E SOUSA, QUE O SUCCEDEU NO VICE-REINADO.* () [word count] [A00_0851].⁴

5. O DHPB integra o conjunto de obras de referência do Grande Houaiss, conforme o informado pelo coordenador do Instituto Antonio Houaiss, Mauro Villar em carta a mim enviada, onde ressalta a revisão de datas que pôde ser realizada no Dicionário Houaiss com base em registros no DHPB e em seu banco de textos;

⁴ **Fonte:** DHPB. Disponível em: <https://dicionarios.fclar.unesp.br/dhpb/> Acesso em: 18 set. 2023.

6. *Produção acadêmica de orientandos no Programa de Pós-graduação em Linguística e Língua Portuguesa da FCLAr/UNESP.* Sob a minha orientação foram defendidas uma dissertação de Mestrado (Domladovac-Silva, 2017) e cinco teses de Doutorado (Peruchi, 2014; Jeronimo, 2018; Bernardo, 2020; Almeida, 2022; Domladovac-Silva, 2022). Foram seis orientações que tiveram como fonte de pesquisa o banco de textos e o próprio DHPB. Dos cinco doutorados, três foram indicados pelo Programa de Pós-Graduação em Linguística e Língua Portuguesa, da FCLAr/UNESP para concorrerem ao Prêmio Capes de Teses. Muito embora dois deles não tenham sido contemplados na premiação final, o terceiro foi indicado neste ano de 2023 e ainda não se tem o resultado. A tese, com título “O Vocabulário da escravidão no banco de dados do ‘Dicionário Histórico do Português do Brasil – séculos XVI, XVII e XVIII’ – conexões léxico-culturais”, defendida por Mayara Aparecida Ribeiro de Almeida em 2022, abriu as portas para que a candidata passasse a integrar a equipe do Instituto Antonio Houaiss, dedicando-se à pesquisa de retrodatações.

CONSIDERAÇÕES FINAIS

A “viagem” realizada de 2005 a 2021 foi longa, repleta de percalços, de momentos bons e de alegria, mas também de muita tristeza e apreensão, mas a chegada a bom porto, como dizem os navegadores, foi de muito êxito, pois, além de ter trazido para a Lexicografia da Língua Portuguesa do Brasil uma contribuição inestimável, permanece como aquela fotografia instantânea de Menéndez Pidal, documentando o passado e oferecendo material linguístico para estudos presentes e futuros.

REFERÊNCIAS

ALMEIDA, Mayara Aparecida Ribeiro de. **Vocabulário da escravidão no banco de dados do “Dicionário Histórico do Português do Brasil - séculos XVI, XVII e XVIII”**: conexões léxico-culturais. 2022. 337 f. Tese (Doutorado em Linguística e Língua Portuguesa) – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências e Letras, Araraquara, 2022.

BERNARDO, Jozimar Luciovanio. **Vocabulário têxtil na língua portuguesa do Brasil Colônia**: tessituras histórico-linguísticas. 2020. 385 f. Tese (Doutorado em Linguística e Língua Portuguesa) – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências e Letras, Araraquara, 2020.

BIDERMAN, Maria Tereza Camargo; MURAKAWA, Clotilde de Almeida Azevedo (org.). **Dicionário Histórico do Português do Brasil**: séculos XVI, XVII e XVIII. Araraquara: FCL-UNESP, 2021. Disponível em: <http://dicionarios.fclar.unesp.br>. Acesso em: 08 set. 2023.

DOMLADOVAC-SILVA, Carolina. **O léxico das enfermidades na obra Erário Mineral (1735), de Luís Gomes Ferreira**. 2017. 206 f. Dissertação (Mestrado em Linguística e Língua Portuguesa) – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências e Letras, Araraquara, 2017.

DOMLADOVAC-SILVA, Carolina. **Um estudo léxico-semântico sobre a farmacopeia do Brasil Colonial em documentos do século XVIII**. 190 f. Tese (Doutorado em Linguística e Língua Portuguesa) – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências e Letras, Araraquara, 2022.

JERONIMO, Gabriela Guimarães. **Dos saberes da roça**: entre plantas, memórias e palavras. 2018. 354 f. Tese (Doutorado em Linguística e Língua Portuguesa) - Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências e Letras, Araraquara, 2018.

MENÉNDEZ PIDAL, Ramón. **Estudios de Lingüística**. Madrid: Espasa-Calpe, 1961.

PERUCHI, Rosane Malusá Gonçalves. **O léxico da apicultura e da melinponicultura no Brasil**: limites e flexibilidade para sua fixação em um dicionário terminológico. 2014. 447 f. Tese (Doutorado em Linguística e Língua Portuguesa) – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências e Letras, Araraquara, 2014.

EXPRESSÕES IDIOMÁTICAS EM DICIONÁRIOS: QUESTÕES INICIAIS EM DIREÇÃO À FRASEOGRAFIA PEDAGÓGICA

*Elizabete Aparecida Marques
Jessica dos Santos Paião*

INTRODUÇÃO

As expressões idiomáticas (EI) são unidades lexicais que pertencem ao acervo vocabular do falante de uma determinada língua e, geralmente, são definidas como lexias complexas indecomponíveis, conotativas e cristalizadas em um idioma pela tradição cultural (Xatara, 1998, p. 170).

No âmbito do ensino e aprendizagem de línguas, os traços caracterizadores dessas unidades, como polilexicalidade, significado translato e, em muitos casos, alto grau de opacidade, poderiam explicar o desafio para o aprendiz de espanhol como língua estrangeira (ELE) em compreender o significado de expressões como *salir rana* (“defraudar”) e *alzar el gallo* (“manifestar soberbia ou arrogância”), pois seu sentido global não resulta da soma dos significados dos itens lexicais isolados que integram a combinatória fixa. O desafio não se restringe, porém, apenas aos discentes. Os professores de línguas, além da questão relacionada ao (re)conhecimento das características intrínsecas às EI, enfrentam uma lacuna de pesquisas que contemplem

aspectos didáticos como o tipo de unidade fraseológica¹ (UF) que deve ser ensinada e sua adequação por nível de aprendizagem. Para complicar a situação, o docente ainda precisa lidar com a escassez de materiais didáticos específicos para o ensino de uma parcela do léxico tão representativa do vocabulário de um povo.

Em face da atual lacuna da fraseodidática em nosso país, consideramos que o dicionário escolar pode ser um aliado indispensável em sala de aula para o ensino e aprendizagem dos diversos tipos de fraseologismos, especialmente as EI. Mas, como as EI são tratadas nesse tipo de obra lexicográfica? A partir desse questionamento inicial, com base nos pressupostos da Lexicografia Pedagógica e da Fraseografia, analisamos o tratamento de expressões idiomáticas em dicionários considerados como de uso pedagógico, tanto de língua portuguesa como de língua espanhola.

O *corpus* de análise está constituído por quatro obras lexicográficas de carácter escolar: *Caldas Aulete dicionário escolar da língua portuguesa* (2012), *Diccionario para la enseñanza de la lengua española para brasileños* (2008), *Diccionario escolar português/español - español/português* (2011) e *Dicionário Larousse Espanhol/Português - Português/Espanhol-Avançado* (2006).

A partir da análise, observamos que quando a expressão idiomática é repertoriada em dicionários cujo público-alvo é o aprendiz de língua (materna ou estrangeira), novas práticas precisam ser adotadas. É nessa perspectiva que chamamos a atenção para a importância de uma vertente pedagógica da Fraseografia, denominada neste trabalho de Fraseografia Pedagógica².

¹ Termo genérico usado indistintamente, neste trabalho, como fraseologismo. *Corpas Pastor* (1996, p. 20, tradução nossa) define as unidades fraseológicas como “unidades léxicas formadas por mais de duas palavras gráficas em seu limite inferior, cujo limite superior se situa no nível de oração composta”.

² Molina Garcia (2006) e Paião (2021) foram, até o momento, os autores que chamaram

Com base nos pressupostos teóricos da Fraseodidática, da Lexicografia Pedagógica e da Fraseografia, analisamos a macro e a microestrutura dos dicionários utilizados como fonte de pesquisa, com o foco principal nos seguintes aspectos: i) proposta lexicográfica, especialmente em busca de informações relativas às EI; ii) estruturação do verbete a fim de verificar a lematização e a descrição das expressões.

1. FRASEODIDÁTICA: algumas considerações

A Fraseodidática (ou didática da fraseologia) é a responsável, tanto em língua materna como em língua estrangeira, pelas discussões quanto às habilidades e às dificuldades relacionadas ao ensino das unidades fraseológicas e, por extensão, à aprendizagem dos fraseologismos. Para Ettinger (2008, p. 96), um dos precursores das pesquisas realizadas sobre o assunto, a Fraseodidática pode ser definida como

[...] una disciplina relativamente nova, ocúpase do ensino e aprendizaxe sistemáticos e con base científica de fraseoloxismos no ensino de idiomas. A súa misión consiste en que os fraseoloxismos se recoñezan, aprendan e empreguen como unidades poliléxicas con significado³.

De modo similar, para González Rey (2010, p. 3), “la phraséodidactique est une discipline très jeune (d’à peine une vingtaine d’années) et très peu connue”⁴. Conforme a autora, além dos esforços

a atenção para a necessidade do desenvolvimento de pesquisas vinculadas a essa área.

³ “uma disciplina relativamente nova, que se ocupa do ensino e aprendizagem sistemáticos e com base científica de fraseologismos no ensino de idiomas. A sua missão consiste em que os fraseologismos sejam reconhecidos, aprendidos e empregados como unidades poliléxicas significativas” (Ettinger, 2008, p. 96, tradução nossa).

⁴ “a fraseodidática é uma disciplina muito jovem (apenas 20 anos) e muito pouco conhecida” (González Rey, 2010, p. 3, tradução nossa).

de Ettinger (2008), a Fraseodidática nasce a partir de estudos realizados por Peter Khün, Regina Hessky, Heinz-Helmut Lüger, entre outros. Visando ao estabelecimento de um modelo didático, essa disciplina reivindica uma função no centro da didática de línguas estrangeiras, uma vez que a presença da fraseologia nesse cenário é quase inexistente (González Rey, 2010). Nesse sentido, seu objetivo é levar os aprendizes a reconhecer, a aprender e a utilizar unidades fraseológicas em situações comunicativas diversas. Ettinger (2008, p. 96) salienta que uma das finalidades da disciplina é “descubrir as intencions particulares do emisor. O aprendiz debe saber por que um falante emprega fraseoloxismos e non simplemente as paráfrases destes”⁵.

A Fraseodidática conjuga um papel fundamental no âmbito do ensino e aprendizagem das unidades fraseológicas, já que além de conhecer e dominar o código da língua, o que Chomsky (1965) chamou de competência linguística, o aprendiz precisa compreender as circunstâncias que perpassam o ato de comunicação, não restringindo-se apenas ao código linguístico, ou seja, é necessário alcançar a competência comunicativa (Hymes, 1972).

Sobre essa última competência, com o intuito de aplicá-la ao ensino de línguas, Canale e Swain (1980 *apud* Conca, 2005) postulam que ela está integrada por quatro componentes: gramatical, discursivo, sociolinguístico e estratégico. Incorporando o componente cultural a esse modelo, Conca (2005) passa a denominá-la como competência fraseológica que, para a autora, compreende tanto a competência linguística como a comunicativa. Nesse sentido, conforme exposto no Quadro 1, podemos verificar que há na competência fraseológica os cinco componentes:

⁵ “Descobrir as intencões particulares do emissor. O aluno deve saber por que um falante usa fraseologismos e não simplesmente suas paráfrases” (Ettinger, 2008, p. 96, tradução nossa).

Quadro 1 – Componentes da competência fraseológica

1. Componente gramatical	As UF possuem regras de combinação, de inserção e de pronúncia.
2. Componente discursivo	As UF apresentam coesão sintática, léxica ou pragmática e são importantes para a coerência semântica.
3. Componente sociolinguístico	As UF funcionam em determinados contextos, são instrumentos que podem identificar grupos e podem ajudar na socialização.
4. Componente estratégico	Ainda que sejam formas simples de expressão, as UF são muito significativas.
5. Componente cultural	O conhecimento fraseológico de uma língua é um aspecto potente de integração cultural.

Fonte: elaborado pelas autoras com base em Conca (2005, p. 78-79)

Para a autora, a competência fraseológica é dinâmica e integra a competência comunicativa, tornando-se fundamental, portanto, a incorporação das unidades fraseológicas nos currículos de todos os níveis escolares do ensino de línguas. Em 1999, no livro *La enseñanza de las unidades fraseológicas*, Penadés Martínez já chamava a atenção para a necessidade de que o aluno aprenda fraseologismos desde o primeiro contato com a língua estrangeira. No entanto, no Brasil, documentos norteadores do ensino, como os Parâmetros Curriculares Nacionais (Brasil, 2000) e Base Nacional Comum Curricular (Brasil, 2018) não fazem referência ao ensino das UF.

Dito isso, quais seriam as vantagens de incorporar as unidades fraseológicas no ensino de línguas? Para essa resposta, recorreremos aos trabalhos de Ettinger (2008) e de González Rey (2010). De acordo com eles, existem dois grupos distintos de estudiosos de didática: a) os fraseófobos (*phraséophobes*), aqueles que acreditam que a abordagem desse tipo de unidade em sala de aula pode ser supérflua e b) os fraseófilos (*phraseophiles*), para quem a aquisição de UF é essencial.

Concordamos com a visão dos fraseófilos, pois acreditamos que as UF são unidades linguísticas e culturais com extrema relevância para o desenvolvimento da competência comunicativa dos estudantes. Tendo isso em vista, vale a pena resgatar, ainda que de forma resumida, os argumentos levantados por González Rey (2010) a favor da didática da fraseologia. São eles: a) desconhecer essas unidades pode causar mal-entendidos; b) as UF são úteis para que os aprendizes de uma língua estrangeira possam expressar-se de modo natural; c) existem incontáveis UF permeando os discursos; d) o ensino de UF pode proporcionar o conhecimento cultural do novo idioma.

De modo a construir uma abordagem voltada para o ensino de unidades fraseológicas, é importante dispor de uma metodologia adequada e de materiais didáticos coerentes e consistentes, sempre levando em consideração que o principal objetivo dessa incorporação é a formação de falantes competentes, tanto em língua materna quanto em língua estrangeira, que sejam capazes de codificar e decodificar as UF nos mais variados contextos de uso. Sobre os materiais didáticos específicos, Penadés Martínez (1999) destaca a importância de obras como livros e dicionários. Nesse cenário, a Fraseologia, a Lexicografia, a Fraseografia, a Fraseodidática e a Lexicografia Pedagógica (doravante LEXPED) estão em estreita relação.

2. LEXICOGRAFIA PEDAGÓGICA E FRASEOGRAFIA

A LEXPED pode ser entendida como “an interdisciplinary problem area that lies roughly half-way between lexicography and language teaching”⁶ (Hartmann, 1999, p. 5). O seu surgimento está

⁶ “Uma área problemática interdisciplinar que fica aproximadamente a meio caminho entre a lexicografia e o ensino de línguas” (Hartmann, 1999, p. 5, tradução nossa).

atrelado aos avanços pedagógicos ocorridos no século XX, um cenário que revolucionou as tendências didáticas sobre o ensino de línguas e que, conseqüentemente, contribuiu para uma mudança na elaboração de dicionários, com vistas a facilitar o processo de aprendizagem, sobretudo, de língua estrangeira. M. West, H. E. Palmer e A. S. Hornby, professores de inglês como língua estrangeira, são os precursores desse movimento (Molina García, 2006).

Para Molina García (2006), existem três aspectos fundamentais da Lexicografia Pedagógica que devem ser levados em consideração. O primeiro deles é a qualidade e adequação das informações à necessidade do possível usuário. Assim como ocorre com qualquer ato de comunicação, é necessário que o lexicógrafo (o emissor nesse caso), ao entrar em contato com o usuário do dicionário (o receptor) mediante o código linguístico, ou seja, as definições utilizadas na obra, os dados gramaticais etc., faça-o de maneira compreensível, pois, se o usuário não for capaz de decodificar os elementos expostos no dicionário, a obra não cumpre o seu objetivo comunicativo. Nesse contexto, além da clareza na comunicação, é preciso levar em consideração o nível de conhecimento que o usuário possui da língua estudada.

O segundo aspecto diz respeito às inovações estruturais e de conteúdo lexicográfico. Para além dos princípios fundamentais da LEXPED, como definições mais claras, exemplos de uso ou abonações adequados à faixa etária dos alunos, descrição gramatical e de uso das UF, os quais colaboram para o acesso aos conhecimentos por parte do aprendiz de língua estrangeira, podemos citar outros aspectos, como a facilidade para encontrar as informações desejadas, tanto na macro quanto na microestrutura das obras. A publicação de dicionários onomasiológicos destinados ao ensino, consoante a Molina García (2006, p. 19), pode ser considerada uma inovação, pois essas obras podem servir de guia para o aprendiz encontrar o significado a partir de sua

forma linguística. Ressaltamos, no entanto, que nem todas as explicações dispostas nos dicionários onomasiológicos podem ser úteis para um estudante, além do mais, muitos repertórios desse tipo não oferecem elementos básicos ao usuário, como definição, indicação gramatical e exemplo de uso.

Os outros avanços dizem respeito ao aperfeiçoamento da apresentação gramatical, do conteúdo semântico (vocabulário orientado ao público-alvo e uso de *corpora* para extrair exemplos de uso), das referências cruzadas e das informações pragmáticas. Essas mudanças contribuem para que o dicionário do aprendiz seja diferente daquele destinado ao falante nativo. Para Kirkpatrick (1985 *apud* Molina García, 2006, p. 20), existem sete características distintivas entre esses dois tipos de obras, a saber: 1) realce na transcrição fonética; 2) vocabulário limitado para as definições; 3) clareza nas acepções; 4) explicação sobre os diferentes significados; 5) tratamento mais exaustivo para as unidades funcionais; 6) uso de exemplos; 7) etiquetas de registro.

O terceiro aspecto fundamental da Lexicografia Pedagógica está relacionado com a prática lexicográfica e o trabalho docente. O professor, a partir de sua compreensão sobre a relevância dos dicionários em sala de aula, deve instruir os aprendizes a manejar esse tipo de repertório e deve evidenciar o potencial informativo do dicionário.

Na mesma direção, Krieger (2011, p. 106) aponta que os princípios que norteiam a LEXPED são: i) a adequação do dicionário ao usuário; ii) o uso produtivo pensado no ensino e aprendizagem de línguas, e iii) a compreensão de que esse tipo de obra é um texto com informações linguísticas, culturais e pragmáticas. Ainda, segundo a autora,

[...] daí resulta seu exponencial papel pedagógico, bem como o princípio de que assim como há livros didáticos adequados aos diferentes níveis de ensino,

de igual modo, deve-se proceder à escolha do dicionário adequado à necessidade de aprendizagem dos alunos (Krieger, 2011, p. 106).

Motivada pela preocupação com a qualidade dos dicionários usados no ensino de línguas, seu principal objetivo está centrado na constituição e no estudo das peculiaridades relacionadas aos dicionários pedagógicos. Nesse sentido, em consonância com Welker (2008), podemos atribuir à LEXPED uma parte teórica e uma prática: a primeira visa a contribuir para o estudo de dicionários voltados para o auxílio na aprendizagem de línguas (estrangeira ou materna), enquanto a segunda objetiva a elaboração de dicionários com função pedagógica.

3. CRITÉRIOS QUE ORIENTARAM A ANÁLISE

Enquanto os dicionários comuns oferecem informações de forma pouco didática (linguagem não acessível ao nível escolar, definição não funcional, por exemplo), os dicionários pedagógicos levam em conta as necessidades e as dificuldades do aprendiz de línguas, ou seja, as escolhas realizadas desde a seleção das unidades lexicais até a estrutura do verbete devem ser pensadas no potencial consulente. E como as EI são tratadas nos dicionários pedagógicos? A fim de responder esse questionamento, com base nos pressupostos da Lexicografia Pedagógica e da Fraseografia, analisamos o tratamento de expressões idiomáticas nos dicionários *Caldas Aulete* (2012), *Señas* (2008), *Diccionario Escolar* (2011) e o *Larousse* (2006), todos voltados para o ensino.

Os critérios utilizados na análise da macro e da microestrutura dos dicionários pesquisados estão enumerados a seguir:

- 1- Estrutura da parte introdutória e apêndices;
- 2- Aspecto teórico da Fraseologia;
- 3- Seleção e quantidade de unidades simples e de unidades complexas registradas;
- 4- Público-alvo da obra;
- 5- Ordenação;
- 6- Localização das EI no verbete;
- 7- Lematização das EI e elementos do contorno;
- 8- Informação gramatical;
- 9- Marcas de uso;
- 10- Definição e a linguagem utilizada no texto definitório;
- 11- Exemplos de uso;
- 12- Equivalentes.

4. ANÁLISE DO TRATAMENTO LEXICOGRÁFICO DAS EI NO *CORPUS* DE ESTUDO

Com base nas obras analisadas, percebemos que apesar de serem considerados dicionários voltados para o ensino, nem todos os princípios da Lexicografia Pedagógica e da Fraseografia foram seguidos na compilação e tratamento das EI. Em primeiro lugar, a falta de clareza nas partes introdutórias para definir o que está sendo considerado como expressão idiomática bem como a não exatidão no número de unidades compiladas saltam à vista.

A seleção do caudal fraseológico frente às unidades lexicais simples é um grande desafio enfrentado pela Lexicografia. No caso

dos dicionários analisados, ainda que a maioria deles não ofereça dados concretos sobre a quantidade de EI registradas, de forma geral, é notável a discrepância entre o número de unidades lexicais simples em relação às complexas que pode ser confirmada no *Caldas Aulete* (2012) que explicita o registro de um total de 1.924 locuções frente a 29.431 unidades simples. Ainda sobre o critério de seleção das EI, com exceção do *Larousse* (2006) que não faz menção a ele, é informado que todas as unidades foram selecionadas levando em consideração a frequência de uso. No entanto, não nos pareceu prudente que o *Señas* (2008), o *Diccionario Escolar* (2011) e o *Larousse* (2006) deixassem de repertoriar unidades como: *hincar los codos*, muito utilizada no contexto escolar, ou *como el perro y el gato*, unidade bastante difundida por falantes da língua espanhola. Também nos chamou atenção que apenas o *Larousse* (2006) tenha repertoriado a unidade *hablar por los codos*, outra expressão de grande profusão entre os falantes da língua. Nessa mesma direção, o *Caldas Aulete* (2012), mesmo afirmando que as locuções selecionadas são as mais usadas na linguagem corrente, optou pela EI *gato escaldado*, em detrimento da EI, *balaio de gatos*, ainda que essa última pareça ser mais recorrente.

Sobre o público-alvo, o *Señas* (2008) destina a obra, de forma geral, a estudantes brasileiros. O *Diccionario Escolar* (2011) informa que os consulentes são estudantes lusófonos de espanhol como língua estrangeira. Já o *Caldas Aulete* (2012) e o *Larousse* (2006) delimitam como possíveis usuários os estudantes a partir do 6º ano do Ensino Fundamental e estudantes em nível avançado de aprendizagem, respectivamente.

No que se refere à ordenação, todas as obras seguem a ordem semasiológica, uma vez que partem da forma para o conteúdo, referendando, assim, uma tendência da lexicografia nos últimos anos.

Os dicionários gerais monolíngues e também os bilíngues, geralmente, não destinam uma entrada específica para as unidades fraseológicas, que são registradas na microestrutura, como uma sub-entrada da cabeça do verbete. Esse procedimento não se difere nos dicionários em análise, pois nas quatro obras as EI são registradas no interior do verbete, como subentrada, após as acepções da entrada principal, e são lematizadas de acordo com a palavra-chave escolhida pelo lexicógrafo. Apenas o dicionário *Señas* (2008) esclarece os critérios de escolha e, conforme pudemos verificar, são dois: i) se uma locução for constituída por elementos da mesma classe de palavras, ela estará incluída na entrada correspondente ao primeiro elemento; e ii) caso contrário, ela estará incluída na entrada do elemento que pertença à classe de palavras mais relevante, a saber: substantivo, verbo, adjetivo, advérbio e preposições. No *Caldas Aulete* (2012) e no *Señas* (2008), as EI aparecem precedidas por um símbolo. Além disso, nos quatro dicionários analisados verificamos que elas são grafadas em negrito e o sinal ~ substitui a palavra-chave. É o caso, por exemplo, do verbete *dente* no *Caldas Aulete*: ■■ ~ **de coelho. Falar entre os ~s**. Ainda que a utilização desse sinal seja um critério adotado pela tradição lexicográfica, é preciso que haja uma reflexão na elaboração de dicionários destinados a aprendizes de língua (estrangeira ou materna) sobre a utilização de sinais e sobre a omissão de vocábulos das expressões idiomáticas. Será que esse método é realmente eficaz para um potencial consulente que esteja cursando o 6º ano do Ensino Fundamental, como é o caso do destinatário do *Caldas Aulete*? Questões como essa devem ser levadas em consideração pelo lexicógrafo e pelo fraseógrafo.

Quanto ao registro de informação gramatical e de elementos do contorno das EI, nenhum dos dicionários analisados contempla esses aspectos, o que se configura como uma lacuna dessas obras de

caráter escolar. Nesse particular, cabe ressaltar que esses tipos de conhecimento linguístico são importantes, pois contribuem para dirimir dúvidas do consulente acerca da classe gramatical da EI e também sobre as restrições de uso (de caráter sintático ou semântico), geralmente registradas entre colchetes, que conformam os elementos do contorno.

As marcas de uso são recursos lexicográficos para indicar variação diacrônica, diatópica, diastrática, diafásica, entre outras, das unidades lexicais que integram a nomenclatura do dicionário. O registro de marcas de uso na microestrutura de dicionários é de fundamental importância, pois oferece indicações acerca do emprego de suas unidades lexicográficas, dentre elas as fraseológicas, em termos de localização espacial e temporal, além de seus níveis sociais e de registro. Nas obras analisadas, observamos, no geral, uma falta de sistematização no registro de marcas de uso. Para algumas EI, os dicionários registram marcas de uso, enquanto para outras não.

Considerando seu caráter pedagógico, espera-se que todos os dicionários analisados apresentem definições claras em linguagem acessível ao público-alvo objetivado. No entanto, apenas o *Caldas Aulete* (2012) e o *Señas* (2008) dispõem de definições para as EI. As duas obras utilizam-se de tipos sinonímicos e parafrásticos de definição, e empregam uma linguagem compreensível e adequada ao nível de ensino do consulente a que se destinam.

A contextualização, por meio de abonações e exemplos de uso, também é de fundamental importância nos dicionários pedagógicos, pois auxilia o aprendiz na compreensão do uso de uma unidade lexical. No caso dos dicionários examinados neste trabalho, nem todos adotam essa prática para as expressões idiomáticas. Assim, todas as EI registradas no *Señas* (2008) possuem exemplificação logo após o

texto da definição. No *Caldas Aulete* (2012), não há sistematização de exemplos, já que para algumas unidades o dicionário oferece contexto de uso enquanto para outras não se utiliza o mesmo procedimento. O *Larousse* (2006) e o *Diccionario escolar* (2011) não fornecem exemplos de uso, prática que, em nossa opinião, compromete a compreensão das EI devido à falta de contextualização que permitiria ao aprendiz não só entender melhor seu sentido, como também observar os empregos possíveis dessas unidades fraseológicas.

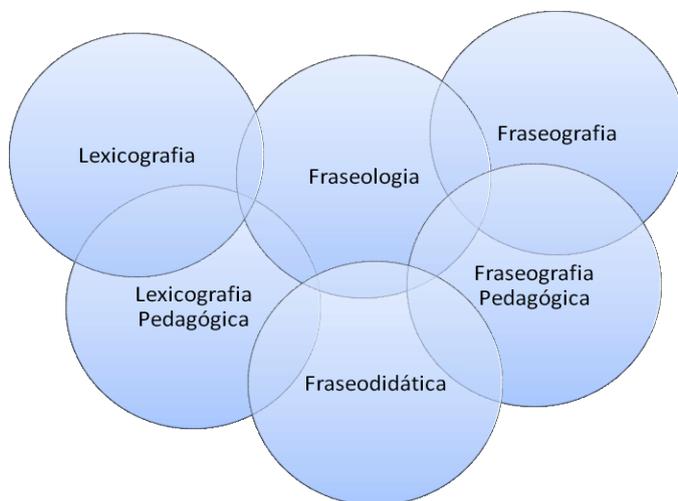
Uma das características da lexicografia bilíngue é o registro de equivalência ou correspondência para o vocabulário compilado. Nesse sentido, os dicionários bilíngues examinados oferecem equivalentes para as EI registradas em sua microestrutura. Em alguns casos, os dicionários proporcionam unidades lexicais simples como equivalentes da expressão. Em outros, a obra oferece uma expressão idiomática correspondente na outra língua, a qual cumpre o papel de equivalente fraseológico.

Finalizando, o exame dos quatro dicionários voltados para o ensino evidencia, em linhas gerais, que as expressões idiomáticas, unidades que fazem parte do patrimônio vocabular de uma comunidade linguística, não recebem um tratamento lexicográfico sistematizado e adequado para os aprendizes de língua, seja estrangeira (espanhol) ou materna, conforme se observa no quadro 2 (anexo), que reúne de forma mais detalhada, a análise descrita e discutida neste tópico. Essa lacuna, no que tange às unidades fraseológicas, está longe de uma prática que compreende o dicionário pedagógico como um tipo de texto com informações linguísticas, culturais e pragmáticas.

Desse modo, a inclusão de unidades fraseológicas em dicionários pedagógicos exige novas práticas, que conjuguem pressupostos teóricos-metodológicos de diversos saberes como a Fraseologia, a

Lexicografia, a Fraseografia, a Lexicografia Pedagógica, a Fraseodidática e a Fraseografia Pedagógica. Nessa perspectiva, a nossa proposta para a interrelação fraseolexicográfica pode ser observada na Figura 1.

Figura 1 – Relação fraseolexicográfica



Fonte: elaborada pelas autoras

Essa relação implicada pode contribuir para o fortalecimento das bases teórico-metodológicas de uma vertente da Fraseografia que, em consonância com Molina García (2006, p. 354), se denomina Fraseografia Pedagógica. Esta, de modo similar à Lexicografia Pedagógica, tem como tarefa o tratamento teórico e prático das UF em dicionários pedagógicos.

CONSIDERAÇÕES FINAIS

A modo de conclusão, ratificamos a ideia de que a Fraseografia é a responsável pelos estudos e pela elaboração de dicionários de unidades fraseológicas de diferentes categorias (colocações, expressões

idiomáticas, provérbios, dentre outras). Quanto às questões relativas ao ensino e aprendizagem desse tipo de unidade lexical, a Fraseodidática desempenha um papel relevante ao ocupar-se de pesquisas que têm como foco a sala de aula, a elaboração e análise de materiais didáticos, bem como de questões relacionadas com o uso de unidades fraseológicas a partir da análise de *corpora* de aprendizes. Por sua vez, cabe à Lexicografia Pedagógica a compilação e tratamento lexicográfico de repertórios destinados ao ensino e aprendizagem de línguas.

Sem desmerecer as obras pesquisadas que, por sinal, são de grande valor e trazem uma contribuição relevante para a Lexicografia e o ensino de línguas, a análise dos dicionários selecionados para este estudo revelou, em linhas gerais, que as unidades fraseológicas, especificamente as EI, foco deste trabalho, não recebem um tratamento lexicográfico sistematizado e adequado para os aprendizes de língua, seja estrangeira (espanhol) ou materna. Dessa maneira, o tratamento lexicográfico de unidades fraseológicas em dicionários pedagógicos implica uma prática que conjugue pressupostos teóricos-metodológicos de diversos saberes, e que coloque no centro de suas preocupações os aprendizes em seu contexto de ensino e aprendizagem de línguas.

REFERÊNCIAS

AULETE, Caldas. **Caldas Aulete**: dicionário escolar da língua portuguesa. Organização de Paulo Geiger. Rio de Janeiro: Lexikon, 2012.

BRASIL. Ministério da Educação. **Base Nacional Comum Curricular**. Brasília: MEC, 2018.

BRASIL. Ministério da Educação. **Parâmetros Curriculares Nacionais (Ensino Médio)**. Parte I - Bases Legais, Parte II - Linguagens, Códigos e suas Tecnologias, Parte III - Ciências da Natureza, Códigos e suas Tecnologias, Parte IV - Ciências Humanas e suas Tecnologias. Brasília: MEC, 2000.

CHOMSKY, Noam. **Aspects of the theory of syntax**. Cambridge: Massachusetts Institute of Technology Press, 1965.

CONCA, Maria. Achegas da fraseoloxía á didáctica da lingua e a literatura. **Cadernos de Fraseoloxía Galega**, Santiago de Compostela, n. 7, p. 75-89, 2005.

CORPAS PASTOR, Gloria. **Manual de fraseología española**. Madrid: Gredos, 1996.

DICCIONARIO ESCOLAR: português-espanhol/ espanhol- português. 3. ed. Madrid: SGEL, 2011.

DICIONÁRIO LAROUSSE: espanhol/português, português/espanhol avançado. 1. ed. São Paulo: Larousse do Brasil, 2006.

ETTINGER, Stefan. Alcances e límites da fraseodidáctica: dez preguntas clave sobre o estado actual da investigación. **Cadernos de Fraseoloxía Galega**, Santiago de Compostela, v. 10, p. 95-127, 2008.

GONZÁLEZ REY, Maribel. **La phraséodidactique en action**: les expressions figées comme objet d'enseignement. La Clé des Langues, 2010. Disponível em: <http://cle.ens-lyon.fr/espanhol/langue/traduction/la-phraséodidactique-en-action-les-expressions-figees-comme-objet-d-enseignement>. Acesso em: 1 fev. 2021.

HARTMANN, Reinhard Rudolf Karl. Lexical Reference Books. What are the Issues? **International Journal of Lexicography**, v. 12, p. 5-11, 1999. Disponível em: <https://academic.oup.com/ijl/article-abstract/12/1/5/1001067?redirectedFrom=fulltext>. Acesso em: 5 fev. 2021.

HYMES, Dell Hathaway. On Communicative Competence. In: PRIDE J. B.; HOLMES, J. (ed.). **Sociolinguistics. Selected Readings**. Harmondsworth: Penguin, 1972. p. 269-293.

KRIEGER, Maria da Graça. Questões de lexicografia pedagógica. In: XATARA, Claudia Maria; BEVILACQUA, Cleci; HUMBLÉ, Philippe. (Org.). **Dicionários na teoria e na prática**: como e para quem são feitos. São Paulo: Parábola Editorial, 2011. p. 103-113.

MOLINA GARCÍA, Daniel. **Fraseología bilingüe**: un enfoque lexicográfico-pedagógico. Granada: Comares, 2006.

PAIÃO, Jessica dos Santos. **Dicionário de expressões idiomáticas verbais da língua espanhola**: uma proposta fraseo-onomasiológica. 2021. 190 f.

Tese (Doutorado em Letras) – Universidade Federal de Mato Grosso do Sul, Três Lagoas, 2021.

PENADÉS MARTÍNEZ, Inmaculada. **La enseñanza de las unidades fraseológicas**. Madrid: Arco Libros, 1999.

SEÑAS. **Diccionario para la enseñanza de la lengua española para brasileños**. São Paulo: Martins Fontes, 2008.

XATARA, Claudia Maria. Tipologia das expressões idiomáticas. **Alfa**: Revista de Linguística, São Paulo, n. 42, p. 169-176, 1998.

WELKER, Herbert Andreas. Lexicografia Pedagógica: definições, história, peculiaridades. In: XATARA, Claudia Maria; BEVILACQUA, Cleci; HUMBLÉ, Philippe. **Lexicografia Pedagógica**: pesquisas e perspectivas. Santa Catarina: NUT, 2008. p. 9-45.

ANEXO

Quadro 2 – Análise dos dicionários voltados para o ensino

	Estrutura da parte introdutória e dos apêndices
CALDAS AULETE	Prefácio (Evanildo Bechara); conceitos que regem a organização do dicionário e da proposta lexicográfica; guia de uso; lista de abreviaturas; paradigmas de conjugação verbal; apêndice gramatical; lista de países e conhecimentos geográficos gerais.
SEÑAS	Apresentação da edição brasileira e da espanhola; características gerais da obra; guia de uso, lista de abreviaturas; fonética da língua espanhola; apêndice gramatical; lista de definidores e glossário português-espanhol.
DICCIONARIO ESCOLAR	Introdução; listas de siglas; características e organização dos vocábulos; parte português-espanhol, características da língua espanhola; parte espanhol-português.
LAROUSSE	Apresentação; abreviaturas; transcrição fonética; parte espanhol-português; guia de comunicação em espanhol; questões relacionadas à geografia; parte português-espanhol; gramática espanhola; países e regiões e paradigma verbal.

	Público-alvo
CALDAS AULETE	Estudantes a partir do 6º ano do Ensino Fundamental.
SEÑAS	Conforme a apresentação da edição brasileira, o dicionário é destinado a estudantes brasileiros. Na apresentação da edição espanhola, que faz referência ao dicionário que serviu de base para o <i>Diccionario Señas</i> , há indicação de que o público-alvo é o estudante de espanhol nível intermediário.
DICCIONARIO ESCOLAR	Estudante lusófono de espanhol como língua estrangeira.
LAROUSSE	Estudantes em nível avançado de aprendizagem.
	Quantidade de unidades simples lematizadas e de unidades complexas
CALDAS AULETE	Na parte introdutória da obra é informado que foram lematizadas 29.431 unidades simples e 1.924 locuções.
SEÑAS	Na parte introdutória da obra é informado que há 22.000, no entanto, não há informação sobre a quantidade de UFs.
DICCIONARIO ESCOLAR	Na parte introdutória da obra é informado que há 100 mil palavras e locuções. Porém, não há informação específica sobre a quantidade de UFs.
LAROUSSE	Na parte introdutória da obra é informado que há 90 mil palavras e expressões, mas não há informação específica sobre a quantidade de UFs.
	Aspecto teórico da Fraseologia
CALDAS AULETE	Locução ou expressão idiomática foram definidas dessa maneira: “uma expressão ou locução em que vocábulos assumem, naquele contexto, um sentido diferente daquele que normalmente têm, constitui uma unidade de significado, ou seja, uma unidade léxica”.
SEÑAS	Na introdução, tem-se a seguinte definição: fraseologia (locução ou expressão), considerando-se locuções as sequências de palavras em que não se pode modificar a ordem de seus componentes.
DICCIONARIO ESCOLAR	Não há referência.
LAROUSSE	Não há referência.

	Ordenação da obra
CALDAS AULETE	Semasiológica
SEÑAS	Semasiológica
DICCIONARIO ESCOLAR	Semasiológica
LAROUSSE	Semasiológica
	Sobre a seleção das UFs
CALDAS AULETE	De acordo com a parte introdutória do dicionário, as locuções selecionadas são as mais usadas na linguagem corrente.
SEÑAS	De forma geral, dicionários e manuais e o Corpus VOX-Biblograf. Não há menção especificando o processo de seleção das UFs.
DICCIONARIO ESCOLAR	De forma geral, é informado na introdução que são recolhidas palavras usuais e não tão usuais, mas de interesse do estudante, nas línguas portuguesa e espanhola. Entretanto, não há menção especificando esse processo de seleção.
LAROUSSE	Não há referência.
	Sobre a lematização
CALDAS AULETE	As unidades são lematizadas a partir de uma palavra-chave. No entanto, não há menção ao critério de como essas palavras são escolhidas.
SEÑAS	A inclusão das unidades dependerá das formas que a compõem. Se uma locução for constituída por elementos da mesma classe de palavras, ela estará incluída na entrada correspondente ao primeiro elemento. Caso contrário, ela estará incluída na entrada do elemento que pertença à classe de palavras mais relevante, a saber: substantivo, verbo, adjetivo, advérbio, preposições.
DICCIONARIO ESCOLAR	Não há referência.
LAROUSSE	As expressões e locuções devem ser localizadas a partir do primeiro substantivo que as forma. Se a expressão não possui um substantivo, deve-se procurá-la pelo verbo.

Localização das UFs no verbete	
CALDAS AULETE	As UFs seguem-se no verbete após as definições das unidades simples. Elas são precedidas por um sinal ■■, que indica o início da área de locuções. Elas são grafadas em negrito, e o sinal ~ substitui a palavra-chave. Por exemplo no verbete dente: ■■ ~ de coelho. Falar entre os ~s.
SEÑAS	As UFs estão localizadas no interior do verbete. Elas são precedidas por um sinal ■, que indica o início da locução. São grafadas em negrito, e o sinal ~ substitui a palavra-chave. Por exemplo no verbete bala: ■ ~ perdida. ■ como una ~.
DICCIONARIO ESCOLAR	As UFs estão localizadas no interior do verbete. Elas são precedidas pela abreviação Loc. São grafadas em negrito, e o sinal ~ substitui a palavra-chave. Por exemplo no verbete gato: cuatro ~ s. Dar/vender ~ por liebre.
LAROUSSE	No interior de cada verbete, as expressões aparecem em negrito, com o símbolo (◊)- representando a entrada correspondente. São grafadas em negrito, e o sinal ~ substitui a palavra-chave. Por exemplo no verbete vaca: ~ s flacas/gordas.
Informação gramatical nas UFs	
CALDAS AULETE	Não há referência.
SEÑAS	Não há referência.
DICCIONARIO ESCOLAR	Não há referência.
LAROUSSE	Não há referência.
Elementos de contorno	
CALDAS AULETE	Não há referência.
SEÑAS	Não há referência.
DICCIONARIO ESCOLAR	Não há referência.
LAROUSSE	Não há referência.

Marcas de uso	
CALDAS AULETE	Não há sistematização. Para algumas unidades foram utilizadas marcas, e para outras, não.
SEÑAS	Não há sistematização. Para algumas unidades foram utilizadas marcas e para outras não.
DICCIONARIO ESCOLAR	Não há sistematização. Para algumas unidades foram utilizadas marcas, e para outras, não.
LAROUSSE	Não há sistematização. Para algumas unidades foram utilizadas marcas, e para outras, não.
Sobre a definição e a linguagem utilizada	
CALDAS AULETE	Não há menção ao tipo de definição. A partir de nossa análise, constatamos que em muitos casos foram utilizadas definições sinonímicas e parafrásticas. Muitas delas não apresentavam definições coerentes. A linguagem utilizada na redação da definição é adequada e compreensível.
SEÑAS	De forma geral, é informado na introdução da obra que as definições foram construídas a partir das unidades contidas em uma lista formada por cerca de 2.000 definidores. A linguagem utilizada na redação da definição é adequada e compreensível.
DICCIONARIO ESCOLAR	Não há definição, mas sim equivalências.
LAROUSSE	Não há definição, mas sim equivalências.
Exemplos de uso	
CALDAS AULETE	Não há sistematização nos exemplos. Para algumas unidades foram utilizados, e para outras, não.
SEÑAS	Em todas as unidades há exemplos de uso.
DICCIONARIO ESCOLAR	Não há exemplos de uso.
LAROUSSE	Não há exemplos de uso.

	Equivalentes
CALDAS AULETE	Por ser um dicionário monolíngue, não cabe essa informação.
SEÑAS	Em alguns casos, são apresentados equivalentes com lexias simples e, em outros casos, são apresentados equivalentes fraseológicos.
DICCIONARIO ESCOLAR	Em alguns casos, são apresentados equivalentes com lexias simples e, em outros casos, são apresentados equivalentes fraseológicos.
LAROUSSE	Em alguns casos, são apresentados equivalentes com lexias simples e, em outros casos, são apresentados equivalentes fraseológicos.

Fonte: elaborado pelas autoras

EL ENFOQUE SINCRÓNICO PARA EL ESTABLECIMIENTO DE LA HOMONIMIA DESTINADA AL DMHE¹

Renato Rodrigues-Pereira
Odair Luiz Nadin
Marín Tapia Kwiecien

INTRODUCCIÓN

No es difícil encontrar el camino.
Lo difícil es saber andarlo, adecuar las fuerzas
del recorrido, elegir y seleccionar los atajos
(Alochis, 2007, p. 11).

El epígrafe que nos sirve de punto de partida para este texto, en una perspectiva analógica, define perfectamente el recorrido que hemos tomado para el establecimiento de la homonimia destinada al DMHE. En efecto, definir una unidad léxica² (UL) como un hecho

¹ DMHE – *Diccionario monolingüe de formas homónimas en español para aprendices brasileños*. Este diccionario se elabora actualmente en el contexto del proyecto institucional de investigación titulado “Lexicografía Pedagógica: elaboración del diccionario monolingüe de formas homónimas en español para aprendices brasileños”, bajo la coordinación del Prof. Dr. Renato Rodrigues-Pereira, en la Universidade Federal de Mato Grosso do Sul, *Campus* de Três Lagoas/MS. Dicho proyecto cuenta con un equipo de investigadores de diferentes instituciones. Para tener acceso a la página web del DMHE, accese: <https://nel.net.br/dmhe>.

² En el contexto de las ciencias del léxico, se suele comprender como *unidad léxica* el sintagma formado por uno, dos o más significantes que juntos, en determinado con-

homonímico o polisémico no suele ser una tarea tan simple como nos parece cuando leemos algunos textos sobre el tema. Hay una variedad de posicionamientos teóricos que no convergen, lo que consideramos normal y productivo en el ámbito de la ciencia. Por un lado, tenemos lingüistas que pertenecen a la línea diacrónica y que buscan delimitar los dos hechos lingüísticos, homonimia y polisemia, a partir de aportes de la etimología. Por otro lado, contamos con investigadores que defienden un enfoque sincrónico, cuyos criterios, en términos generales, verifican la divergencia de significados con relación al significante en un determinado estado de la lengua.

En el contexto de la Lexicografía, una ciencia que hoy goza de principios teóricos y metodológicos bien determinados, también podemos recurrir a criterios de naturaleza diacrónica o sincrónica en el proceso de inventariar formas homónimas destinadas a un repertorio lexicográfico. La decisión por cuál fundamento escoger; no obstante, generalmente, es tomada de acuerdo con el tipo de diccionario que, a su vez, es elaborado en conformidad con las necesidades de los potenciales consultantes en sus diversos niveles de competencia en la(s) lengua(s). El hecho es que resulta imprescindible que el lexicógrafo y su equipo tomen decisiones teóricas y metodológicas coherentes con los objetivos relacionados al proyecto lexicográfico que estén elaborando, puesto que varios son los principios epistemológicos que sustentan esta tarea.

En este texto, presentamos el resultado de la segunda etapa³ de la investigación empírica que hemos realizado con el objetivo de

texto de uso, poseen unidad de sentido, de forma que la separación de tales unidades no sería posible sin deshacer el valor de contenido que es prerrogativo. Sobre este asunto, sugerimos la lectura de Biderman (2005), Morante Vallejo (2005), Rodrigues-Pereira; Zacarias; Nadin (2019), por ejemplo. Para la investigación que presentamos en este capítulo, utilizamos unidades léxicas formadas de solo un significante.

³ Los resultados de la primera etapa están divulgados en Pereira (2018) y Rodrigues-

demostrar la viabilidad de considerar el enfoque sincrónico para la definición de Unidades Léxicas Homónimas (ULH) en el proceso de inventario de candidatos a homónimos que compongan la nomenclatura de diccionarios pedagógicos de español como lengua extranjera, en especial para el *DMHE*, en proceso de elaboración.

Para ello, establecemos los siguientes objetivos específicos: i) averiguar en qué medida los informantes diferencian una ULH de una UL polisémica, a partir de la aplicación de una encuesta elaborada a tal fin; y ii) demostrar la pertinencia de criterios de naturaleza sincrónica para la definición de homonimia destinada a diccionarios pedagógicos de lenguas. Por tanto, nos orientamos por principios teóricos y metodológicos de la Semántica y de la Lexicografía Pedagógica (LEXPED).

1. FUNDAMENTACIÓN TEÓRICA⁴

Para el inventario de los candidatos a homónimos del *DMHE*, tuvimos que decidir sobre cuál (es) criterio (s) sustentarían la definición de una UL en un caso de homonimia o de polisemia. Como ya hemos mencionado en la introducción de este texto, el enfoque adoptado fue el sincrónico, cuyos principios se rigen, específicamente, por la divergencia semántica que un significante adquiere en un determinado estado de la lengua. Sin embargo, para justificar

Pereira; Nadin (2020), en los cuales, a partir de datos de la lengua portuguesa, los autores realizaron un estudio piloto con vistas a verificar si sería posible considerar criterios sincrónicos para el establecimiento de la homonimia en diccionarios elaborados en el ámbito de la LEXPED.

⁴ Una primera versión de esta fundamentación teórica fue escrita para el texto “Bases teóricas e metodológicas para o inventario de candidatos a homónimos destinados a diccionarios pedagógicos” (Rodrigues-Pereira; Nadin, 2020). Para este capítulo, presentamos algunas reflexiones de acuerdo con el objetivo de este texto.

nuestra decisión, presentamos, en la secuencia, algunos principios y reflexiones en conformidad con las importantes contribuciones de los diferentes autores visitados y nuestros objetivos.

Ullmann (1964, p. 364-374) presenta tres procesos generadores de homonimia: i) convergencia fonética; ii) divergencia semántica; y iii) influencia extranjera. El primero, por tratarse de un proceso que considera el enfoque diacrónico, procede la verificación de si dos o más palabras, al ser idénticas en un determinado momento de la lengua, fueron, en sus orígenes, diferentes; de forma que si, con ese tipo de investigación, se evidencian tales características, estaremos ante un caso típico de homonimia en una perspectiva diacrónica.

El segundo proceso, el de la divergencia semántica, considera el alejamiento de sentidos que hay entre dos o más palabras en un estado de lengua. Para Ullmann (1964), ese hecho ocurre “cuando dos o más significados de la misma palabra se separan de tal modo que no haya ninguna conexión evidente entre ellas”, así que “la polisemia dará lugar a la homonimia y la unidad de la palabra será destruida” (Ullmann, 1964, p. 368). Para estudiosos que son defensores de otros enfoques, como el diacrónico, por ejemplo Werner (1982) y Porto Dapena (2002), la divergencia semántica resulta en un problema para que se puedan rescatar las varias polisemias de una misma palabra. Nosotros, sin embargo, así como Biderman (1984, 1991, 1998), Zavaglia (2003), Pereira (2018), Rodrigues-Pereira; Nadin (2020), entre otros, defendemos y demostramos que es posible la delimitación de unidades léxicas como casos de homonimia o de polisemia en una perspectiva sincrónica.

Para el tercer proceso, Ullmann (1964, p. 373) afirma que la influencia extranjera es “una forma especial de desarrollos fonéticos convergentes” y que ocurre “cuando una palabra de préstamo se

establece con firmeza en su nuevo ambiente, adaptándose al sistema fonético local, lo que favorece su participación posterior de los cambios normales de sonido”. Además, aclara que la forma de préstamo puede coincidir con otras palabras de la lengua de origen.

Lyons (1977), a su vez, nos presenta algunos criterios de delimitación entre homonimia y polisemia, de forma que el criterio etimológico, los de la maximización de la homonimia y de la polisemia son discutidos por el autor. Específicamente sobre el criterio diacrónico, el investigador esclarece:

Un criterio, hecho explícitamente en la información etimológica anexada en muchas entradas de diccionario, es el conocimiento del lexicólogo sobre la derivación histórica de palabras. Ese criterio generalmente es aceptado como una condición suficiente, aunque no necesaria, que los lexemas en cuestión deban haber desarrollado a partir de lo que eran para lexemas distintos en algunos momentos anteriores de la lengua⁵ (Lyons, 1977, p. 550, traducción nuestra).

En conformidad con este criterio, Lyons (1977) utiliza el ejemplo *porto* – vino y *porto* – ciudad portuguesa para argumentar contra el criterio etimológico, pues, de acuerdo con el autor, “si decimos que *porto* 1 y *porto* 2 son etimológicamente relacionados, entonces, depende de lo alto que estamos preparados para ir, cuando

⁵ One criterion, which is made explicit in the etymological information that is appended to many dictionary entries, is the lexicographer’s knowledge of the historical derivation of words. It is generally taken to be a sufficient, though not a necessary, condition of homonymy that the lexemes in question should be known to have developed from what were for mall y distinct lexemes in some earlier stage of the language (Lyons, 1977, p. 550).

tenemos las evidencias, en trazar la historia de las palabras⁶ (Lyons, 1977, p. 551, traducción nuestra).

Se percibe que trazar la historia de las palabras es un procedimiento de interés a los estudios diacrónicos porque devela mucho sobre procesos morfológicos, así como sobre la historia social, lingüística y cultural de un pueblo. No obstante, las investigaciones en este ámbito no suelen ser muy productivas, en términos cuantitativos, como ya resaltaba Biderman (1991), no forneciendo bases científicas seguras para el establecimiento del étimo de gran númea lengua portuguesa⁷ (Biderman, 1991, p. 287, traducción nuestra). Lo mismo se verifica en Zavaglia (2003, p. 01, traducción nuestra), que resalta que aún es “[...] difícil adoptar como cr para quien “los estudios etimológicos sobre el léxico portugués son pobres y precariositerio b deásico de identificación de un ítem lexical, el estudio ro de las palabras de ldiacrónico⁸”.

De acuerdo con nuestros estudios⁹, percibimos que muchos lingüistas y lexicógrafos acaban no optando por el criterio etimológico al momento de inventariar homónimos para componer la nomenclatura de diccionarios. Generalmente, la alegación más recurrente, basada en los postulados de la sincronía, es la que, para el hablante nativo, es normalmente desconocida la etimología de las palabras que utiliza y su interpretación es indiferente (Lyons, 1997).

⁶ “Whether we say that ‘part1 ‘and ‘port2 · are etymologically related, therefore, depends upon how far we are prepared to go, when we have the evidence, in tracing the history of words” (Lyons, 1977, p. 551).

⁷ “[...] os estudos etimológicos sobre o léxico português são pobres e precários, não fornecendo bases científicas seguras para o estabelecimento do étimo de grande número das palavras da língua portuguesa” (Biderman, 1991, p. 287).

⁸ “[...] difícil adotar como critério básico de identificação de um item lexical, o estudo diacrónico” (Zavaglia, 2003, p. 01).

⁹ Presentados en Pereira (2018).

En síntesis, se puede definir homonimia, desde un punto de vista más general y común, como la relación entre palabras que poseen la misma forma, pero que difieren en el significado, y no solo por el hecho de poseer significados diferentes, sino también por ser completamente extraño el uno al otro.

Werner (1982), en su texto, discurre con respecto a la distinción entre homonimia y polisemia, y el proceso de lematización de esas unidades léxicas en diccionarios, presentándonos algunos criterios para la diferenciación de esos dos fenómenos. Entre ellos, enfatizamos dos: i) el criterio etimológico; y ii) el criterio de la conciencia lingüística de los usuarios.

Para el criterio etimológico, el autor explica que la homonimia se da cuando los diferentes contenidos corresponden a significantes iguales desde que, en sus orígenes, tuvieron significantes diferentes. La polisemia, distintamente, ocurre cuando contenidos diferentes corresponden a significantes iguales desde que, a partir de un punto de vista diacrónico, tienen origen idéntico. Nos encontramos, de ese modo, con casos de convergencia diacrónica en el plano de la expresión y divergencia en el plano del contenido.

En lo que se refiere al criterio ii), el investigador resalta la existencia de la homonimia cuando el hablante no establece ninguna relación entre los diferentes contenidos de una única forma en el plano de la expresión y, de forma opuesta, hay polisemia cuando, en la conciencia del hablante, existe una relación entre los distintos contenidos que pueden corresponder a solo una forma en el plano de la expresión. Vale destacar que, de acuerdo con Werner (1982), este criterio no es pertinente para una distinción científica entre homonimia y polisemia, puesto que no se puede determinar de forma objetiva lo que el hablante de una lengua puede o no establecer como

relación a una determinada unidad lingüística. Pese a esto, como presentamos más adelante, esta comprensión no suele ser común a muchos estudiosos.

Aún de acuerdo con lo que plantea Werner (1982), se propone que se sigan los presupuestos teóricos de la semántica estructural, de forma que exista la identificación de elementos comunes de sememas para los casos de identidad en el plano de la expresión y, diferentemente, la divergencia en el plano del contenido. De ese modo, por un lado, habría homonimia cuando, en el plano del contenido, esos sememas no poseyeran ningún sema en común y, por otro, habría polisemia cuando, a una única forma, correspondieran varios sememas que poseyeran, por lo menos, un semema en común en el plano de la expresión. Por el contrario, tratándose de la elaboración de diccionarios, y en una perspectiva sincrónica, Biderman (1984) discute al respecto de los fenómenos de la homonimia y de la polisemia, señalándolos como algunos de los problemas que interfieren en la elaboración de obras lexicográficas, puesto que influyen en cuestiones relacionadas tanto a la extensión de la macroestructura como en la microestructura de los diccionarios.

Como se puede percibir por las palabras de la investigadora, la tesis de la conciencia del hablante adquiere un lugar destacado en la labor lexicográfica, sobre todo en el contexto de la LEXPED. Al justificar la práctica moderna, la autora explica que:

En la moderna Lexicografía, sobre todo en aquella que se hace en Francia, el procedimiento ha sido considerar homónimas las palabras de grafía idéntica (mismo significante) y significados muy distintos, a punto de ser difícil para el hablante identificar semas comunes a los dos o más

homónimos¹⁰ (Biderman, 1984, p. 143, traducción nuestra).

En esta misma línea, Biderman (1991) hace referencia a la investigación de Silva (1986), que sometió a veinticuatro hablantes del portugués, estudiantes de Humanidades de la Universidad Católica de Braga, a *tests* de identificación de unidades léxicas ambiguas. Para ello, el autor elaboró una serie de frases dispuestas en pares sobre 100 significantes, 13 de las cuales eran clasificadas como homónimas en el diccionario de lengua portuguesa Aurelio. De acuerdo con Silva (1986, *apud* Biderman, 1991), los informantes examinaron los pares de frases y fueron señalando los grados de similitud en una escala de 0 a 4, considerando, pues, la palabra clave en cuestión. Esta escala destacaba:

1) ausencia total de similitud de sentido (0);
2) grados diversos de similitudes semánticas (1 a 4) palabras polisémicas. El resultado de la investigación mostró que hay un alto grado de acuerdo entre los hablantes en la discriminación de estas dos categorías: media del 78 %. Y nunca ocurrió un desacuerdo grande entre los hablantes, es decir: un acuerdo inferior al 50 % Silva (1986, *apud* Biderman, 1991, p. 288).

Biderman (1991) subraya que la investigación realizada por Silva (1986) ha evidenciado que la distinción entre polisemia y homonimia en nivel de la teoría lingüística no solo es posible, sino necesaria. Según Silva (1986, p. 10-11, *apud* Biderman, 1991), “Los criterios teóricos propuestos presentan resultados adecuados a la

¹⁰ Na moderna Lexicografia, sobretudo aquela que se faz na França, o procedimento tem sido considerar homônimas palavras de grafia idêntica (mesmo significante) e significados muito distintos, a ponto de ser difícil para o falante identificar semas comuns aos dois ou mais homônimos (Biderman, 1984, p. 143).

realidad de estos fenómenos y su funcionamiento podrá y deberá ser experimentado por el lexicógrafo”.

Además, como forma de afianzar aún más nuestras reflexiones, buscamos lo que Zavaglia (2002) nos presenta acerca del contexto de uso de una UL y la importancia de una mirada atenta para el nivel discursivo, puesto que, dependiendo del entorno que la palabra está inserida, podemos reconocerla como polisemia o homonimia. Sobre eso, la autora discurre específicamente acerca del sistema y la norma lingüística, presentando la hipótesis de que:

Si la homonimia es un fenómeno en que hay un significante para dos significados radicalmente diferentes, es decir, sin compartición de rasgos comunes, su existencia habrá de ocurrir en el nivel del discurso, pues será el contexto en el cual un homónimo habrá de encontrarse inserido que determinará su realización o como *lexia*¹ o como *lexia*², y así sucesivamente para tantos cuantos serán sus significados. La cuestión que emerge, no obstante, es la siguiente: ¿serán las realizaciones homonímicas de una *lexia* comprobadas por su uso, es decir, por la norma lingüística de una comunidad, o serán posibilidades y libertades que el sistema lingüístico posibilita? [...] ¹¹ (Zavaglia, 2002, p. 98, traducción nuestra).

¹¹ Se a homonímia é o fenómeno em que se tem um significante para dois significados radicalmente diferentes, ou seja, sem compartilhamento de traços comuns, a sua existência dar-se-á no nível do discurso, pois será o contexto no qual um homônimo encontrar-se-á inserido que determinará a sua realização ou como *lexia*¹ ou como *lexia*² e assim sucessivamente para tantos quantos forem seus significados. A questão que surge, porém, é a seguinte: serão as realizações homonímicas de uma *lexia* comprovadas pelo seu uso, ou seja, pela norma linguística de uma comunidade, ou serão possibilidades e libertades que o sistema linguístico possibilita? [...] (Zavaglia, 2002, p. 98).

A partir de lo que presentamos en esta investigación, hemos podido percibir que el contexto parece delimitar algunos casos especiales de homonimia, como el caso de *vino* y *canto*, por ejemplo, presentados en la sección *presentación y análisis de los datos* de este texto.

En ese escenario y de acuerdo con las contribuciones de Biderman (1984; 1991) y de Silva (1986, *apud* Biderman, 1991), presentadas anteriormente, comenzamos nuestras investigaciones empíricas con vistas a verificar de qué modo sería, de hecho, pertinente considerar el enfoque sincrónico en el proceso de inventario de candidatos a homónimos al *DMHE*. Por lo tanto, siguiendo lo planteado por Silva (1986), buscamos experimentar el funcionamiento de su propuesta. En un primer momento, con datos homónimos y polisémicos del portugués (Pereira, 2018; Rodrigues-Pereira; Nadin, 2020) y ahora con datos de la lengua española, como lo presentamos en la secuencia, descripta a continuación.

2. PROCEDIMIENTOS METODOLÓGICOS

Considerando los objetivos establecidos para esta etapa de la investigación, tomamos los siguientes procedimientos: i) elaboración de la encuesta; ii) aplicación de la encuesta¹² a hablantes nativos de lengua española, específicamente a alumnos de la UNC - Universidad Nacional de Córdoba/Argentina; iii) análisis de los datos, con vistas a verificar si lo que fue constatado por Pereira (2018) y Rodrigues-Pereira; Nadin (2020), bajo la investigación empírica piloto realizada, se comprueba también en un contexto hispanohablante.

¹² La encuesta fue aplicada vía *formulario del Google*.

Para i), en un primer momento, escogemos aleatoriamente diez (10) UL que pueden ser consideradas formas homónimas homógrafas, homófonas o polisémicas: **banco, vino, carta, gato, taco, coma, llama, terraza, canto y tienda**. En un segundo momento, organizamos veintiún (21) pares de oraciones. Para tal procedimiento, utilizamos, cuando fuese posible, los ejemplos de uso disponibles en dos diccionarios: el *Diccionario Salamanca – Español para extranjeros* (Gutiérrez Cuadrado; Pascual Rodríguez, 2006); y el *Diccionario de Español para Extranjeros – con el español que se habla hoy en España y en América Latina* (González, 2005).

Los veintiún (21) pares de oraciones fueron reorganizados en diez (10) grupos de acuerdo con las UL elegidas para la investigación, de forma que en la palabra clave 1, *banco*, fueron agrupados todos los pares con la UL *banco*; en la palabra clave 2, *vino*, fueron agrupadas las oraciones referentes a la UL *vino*, y así, por consiguiente, como demostramos a partir de la encuesta que ejemplificamos en la secuencia.

Para cada par, elaboramos dos (2) preguntas: la primera, ***¿Hay relación de sentido entre los términos destacados?***, tuvo el objetivo de verificar si, para el informante, las UL en destaque poseen algún tipo de relación semántica. A partir de esta pregunta, consideramos un caso de homonimia cuando, para la mayoría de los informantes, no hay ningún vínculo semántico entre las UL, lo que difiere de un caso de polisemia cuando, también para la mayoría, hay algún tipo de asociación (Pereira, 2018; Rodrigues-Pereira; Nadin, 2020); ya la segunda pregunta, ***¿cuál es el grado de relación entre los términos destacados?*** Marca el número que mejor representa tu comprensión, objetivó identificar el grado de similitud entre las UL para aquellos informantes que comprendieron haber relación de sentido entre las

lexías en destaque. De ese modo, cero (0) corresponde a la inexistencia y cuatro (4) al grado máximo de relación de sentido entre las UL¹³.

La encuesta titulada “Investigación semántica”, que ejemplificamos a partir de un segmento a continuación, cuenta también con una explicación destinada a los informantes que se dispusieron a participar de la investigación, a saber:

Cuadro 1 – Investigación semántica

Estimado (a):

El cuestionario que se sigue forma parte de una encuesta respecto a una investigación semántica que estamos realizando con el intuito de verificar en qué medida un hablante nativo del español percibe relaciones de sentido entre algunas palabras de la lengua. Por lo tanto, resaltamos que no hay error en sus respuestas; solo lo que entiende sobre la posible relación o no entre las palabras.

Con tus respuestas, podremos continuar una investigación que, a la vez, resultará en el perfeccionamiento de materiales didácticos destinados a aprendices del español como lengua extranjera, puesto que el estudio de los significados de las palabras y los usos en contextos específicos contribuyen con la enseñanza y el aprendizaje de lenguas.

Muchas gracias por tu participación y colaboración.

Un cordial saludo.

I A partir de la lectura de los ejemplos en las frases que se ubican inmediatamente abajo, observa los términos destacados y responde lo que se pide.

¹³ Los análisis relacionados a la segunda pregunta de la encuesta serán divulgados en publicación posterior.

Palabra-clave 1: banco	
<ul style="list-style-type: none"> • En la plaza estuve buscando un BANCO donde sentarme (<i>DSEE</i>, 2006). • Los BANCOS de arena son peligrosos para las embarcaciones (adaptado de <i>DSEE</i>, 2006). 	
¿Hay relación de sentido entre los términos destacados?	Sí () No ()
¿Cuál es el grado de relación de sentido entre los términos destacados? Marca el número que mejor representa tu comprensión.	
0 () 1 () 2 () 3 () 4 ()	
[...] ¹⁴	
Palabra-clave 2: vino	
<ul style="list-style-type: none"> • Una buena comida debe acompañarse con un buen VINO (<i>DEECMG</i>, 2005). • La casa me VINO por herencia (<i>DEECMG</i>, 2005). 	
¿Hay relación de sentido entre los términos destacados?	Sí () No ()
¿Cuál es el grado de relación de sentido entre los términos destacados? Marca el número que mejor representa tu comprensión.	
0 () 1 () 2 () 3 () 4 ()	
[...] ¹⁵	

Fuente: elaboración propia

Para el procedimiento ii), la aplicación de la encuesta fue realizada a un total de ciento ocho (108) informantes. Para ello, establecemos contacto con alumnos de los cursos de Profesorado, Licenciado y Traductorado de Inglés; Profesorado y Licenciado de Español Lengua Materna y Lengua Extranjera; y Profesorado de Portugués, todos de la

¹⁴ Para la UL *banco*, hay más dos (2) combinaciones, como demostramos en la próxima sección de este capítulo.

¹⁵ Todos los otros pares, así como los datos de la investigación, son presentados y discutidos en la próxima sección.

UNC - Universidad Nacional de Córdoba/Argentina. Con el contacto, les explicamos, de forma general, sobre la investigación y nuestra necesidad de que respondieran una encuesta, y que, por eso, se los invitaba a ser informantes del proyecto. Resaltamos, pues, que sería muy importante que participaran, puesto que los resultados nos brindarían información que contribuiría al desarrollo de materiales didácticos que serían útiles en los contextos de enseñanza y de aprendizaje de lenguas. Tras la explicación y su aceptación, les destacamos, a los entonces informantes, que no había respuestas correctas o equivocadas, sino su entendimiento respecto a las palabras destacadas en los pares de oraciones. Además, les explicamos que, después de las respuestas a la encuesta, les explicaríamos detalles sobre la investigación y que no se los haríamos antes para no interferir en sus respuestas.

Presentamos el análisis de los datos en la secuencia de los siguientes párrafos que componen este texto.

3. PRESENTACIÓN Y ANÁLISIS DE LOS RESULTADOS

Para la presentación de todos los pares de oraciones, así como los resultados obtenidos a partir de la aplicación de la encuesta a los ciento ocho (108) informantes, organizamos la Tabla I, ubicada más abajo, con el intento de proporcionar al lector una mirada más objetiva de los datos.

En la primera columna, registramos las UL que fueron destacadas en cada par de oración, a recordar: “Los veintiún (21) pares de oraciones fueron reorganizados en diez (10) grupos de acuerdo con las UL elegidas para la investigación [...]”. En la segunda, disponemos todos los pares de oraciones, de acuerdo con la organización explicada

en la sección de los procedimientos metodológicos *ut supra*. Ya para la tercera y cuarta columnas, registramos los datos porcentuales de respuestas positivas y negativas, de forma que, en una perspectiva sincrónica, consideramos una UL como un caso de homonimia cuando para la mayoría de los informantes no haya tenido ninguna relación de sentido; distintamente de aquellas UL que poseen algún rasgo de sentido también para la mayoría de los informantes que, en este caso, se encuadran en los casos de polisemia (Pereira, 2018; Rodrigues-Pereira; Nadin, 2020).

Tabla 1 – Datos de la investigación semántica

UL ¹⁶	Oraciones	Positivas %	Negativas %
Banco	<p>i)</p> <ul style="list-style-type: none"> • En la plaza estuve buscando un BANCO donde sentarme (DSEE, 2006). • Los BANCOS de arena son peligrosos para las embarcaciones (Adaptado de DSEE, 2006). 	39,8 %	60,2 %
	<p>ii)</p> <ul style="list-style-type: none"> • En la plaza estuve buscando un BANCO donde sentarme (DSEE, 2006). • Fui al BANCO para abrir una cuenta corriente (DEECMG, 2005). 	23,1 %	76,9 %
	<p>iii)</p> <ul style="list-style-type: none"> • En la plaza estuve buscando un BANCO donde sentarme (DSEE, 2006). • No te preocupes si no tienes plata para el cine, yo te BANCO. 	36,1 %	63,9 %
Vino	<ul style="list-style-type: none"> • Una buena comida debe acompañarse con un buen VINO (DEECMG, 2005). • La casa me VINO por herencia (DEECMG, 2005). 	5,6 %	94,4 %

¹⁶ UL – Unidades léxicas

Carta	<p>i)</p> <ul style="list-style-type: none"> • ¿No has leído la CARTA que te envié? (<i>DEECMG</i>, 2005). • En la CARTA pone el precio de cada plato (<i>DEECMG</i>, 2005). 	64,8 %	35,2 %
	<p>ii)</p> <ul style="list-style-type: none"> • Con la ayuda de la CARTA náutica, podemos viajar tranquilos (adaptado de <i>DEECMG</i>, 2005). • La baraja española tiene cuarenta CARTAS (<i>DEECMG</i>, 2005). 	31,5 %	68,5 %
Gato	<p>i)</p> <ul style="list-style-type: none"> • A los GATOS les gusta el pescado (<i>DEECMG</i>, 2005). • Para cambiar la rueda pinchada hay que levantar el auto con el GATO (<i>DEECMG</i>, 2005). 	16,7 %	83,3 %
	<p>ii)</p> <ul style="list-style-type: none"> • A los GATOS les gusta el pescado (<i>DEECMG</i>, 2005). • Fui a tomar unas copas con unos GATOS (<i>DEECMG</i>, 2005). Coloquialismo. Persona que ha nacido en Madrid. 	30,6 %	69,4 %
Taco	<p>i)</p> <ul style="list-style-type: none"> • Mi amigo mexicano prepara unos TACOS de pollo muy ricos (<i>DEECMG</i>, 2005). (tortilla de maíz) • Como no tengo tiempo de ir a comer, me echaré un TACO por acá (<i>DEECMG</i>, 2005). (comida ligera) 	88 %	12 %
	<p>ii)</p> <ul style="list-style-type: none"> • Una hoja del TACO del calendario (<i>DEECMG</i>, 2005). • Como aperitivo tomamos unos TACOS de jamón (<i>DEECMG</i>, 2005). 	6,5 %	93,5 %

Coma	<p>i)</p> <ul style="list-style-type: none"> • El accidentado estuvo un mes en estado de COMA (DSEE, 2006). • Los elementos de una enumeración van separados por una COMA delante y otra, atrás (DSEE, 2006). 	31,5 %	68,5 %
	<p>ii)</p> <ul style="list-style-type: none"> • El accidentado estuvo un mes en estado de COMA (DSEE, 2006). • La madre quiere que el niño COMA frutas, verduras y legumbres. 	1,9 %	98,1 %
Llama	<p>i)</p> <ul style="list-style-type: none"> • Las LLAMAS del incendio arrasaron varias hectáreas (DEECMG, 2005). • Mantenían viva las LLAMAS de su amor (DEECMG, 2005). 	88,9 %	11,1 %
	<p>ii)</p> <ul style="list-style-type: none"> • Las LLAMAS del incendio arrasaron varias hectáreas (DEECMG, 2005). • Las LLAMAS son animales de la familia de los camellos (DEECMG, 2005). 	13,9 %	86,1 %
	<p>iii)</p> <ul style="list-style-type: none"> • La LLAMA es un animal de la familia de los camellos (DEECMG, 2005). • Pedrito, LLAMA a tus hermanos para que tomen la merienda. 	5,6 %	94,4 %
Terraza	<ul style="list-style-type: none"> • Mi casa tiene TERRAZA (DEECMG, 2005). • Nos sentamos en una mesa de la TERRAZA del bar (DEECMG, 2005). 	99,1 %	0,9 %

Canto	<p>i)</p> <ul style="list-style-type: none"> • Creo que está dotada para el CANTO (DEECMG, 2005). • Estudio CANTO en el conservatorio (DEECMG, 2005). 	98,1 %	1,9 %
	<p>ii)</p> <ul style="list-style-type: none"> • Estudio CANTO en el conservatorio (DEECMG, 2005). • Cuidado, no te des con el CANTO de la mesa (DSEE, 2006). 	7,4 %	92,6 %
	<p>iii)</p> <ul style="list-style-type: none"> • Estudio CANTO en el conservatorio (DEECMG, 2005). • Me siento feliz cuando CANTO a mis alumnos. 	97,2 %	2,8 %
Tienda	<p>i)</p> <ul style="list-style-type: none"> • En esta calle hay muchas TIENDAS de zapatos (adaptado de DSEE, 2006). • Los excursionistas llevaban varias TIENDAS para acampar en la montaña (DSEE, 2006). 	38,9 %	61,1 %
	<p>ii)</p> <ul style="list-style-type: none"> • En esta calle hay muchas TIENDAS de zapatos (Adaptado de DSEE, 2006). • Señor Juan, por favor, TIENDA la cama. 	2,8 %	97,2 %

Fuente: Elaboración propia

En el grupo *banco*, para los tres pares de oraciones registrados en la Tabla I, obtuvimos datos que nos posibilitan considerar la ULH *banco* como un caso de homonimia, puesto que, respectivamente, para el 60,2 %, el 76,9 % y el 63,9 % de los informantes no hay relación de sentido entre las palabras destacadas.

Para el primer par, en especial: “En la plaza estuve buscando un **BANCO** donde sentarme” (DSEE, 2006); y “Los **BANCOS** de arena

son peligrosos para las embarcaciones” (adaptado de *DSEE*, 2006), teníamos como hipótesis, así como Pereira (2018) y Rodrigues-Pereira; Nadin (2020), que sería un caso de polisemia por dos motivos: i) el *banco de arena* de un río o de un riachuelo nos permite que caminemos sobre él o incluso que nos asentemos para tomar un baño de sol, como ya lo hicimos en nuestra época de juventud, por ejemplo; ii) la embarcación puede no seguir su ruta cuando se depara con un *banco de arena*, de manera que acaban asentándose en él hasta que el nivel de agua suba o surja alguna ayuda para reflotarla.

Sin embargo, nuestra hipótesis sigue siendo refutada, así como en la investigación piloto ya mencionada, lo que nos lleva a confirmar la necesidad de reaplicar los cuestionarios tanto con datos de la lengua portuguesa, como con datos de la lengua española, pero con públicos distintos y, también, con cambio de orden de los grupos, de forma que el grupo de pares de palabras de la ULH *banco* no sea ya el primero. De ese modo, podremos verificar si son pertinentes los motivos inferidos por Pereira (2018), Rodrigues-Pereira; Nadin (2020)¹⁷ y que continuamos infiriendo en esta investigación que ahora presentamos. Con ello, esperamos sanar cualquier duda, para que no haya ninguna, o, de cualquier manera, el mínimo posible de vacíos respecto a la cientificidad de nuestra investigación.

En la segunda UL, *vinho*, con solo un par de oraciones: “Una buena comida debe acompañarse con un buen **VINO**” (*DEECMG*, 2005) y “La casa me **VINO** por herencia” (*DEECMG*, 2005), obtuvimos

¹⁷ Los motivos presentados por los autores se refieren, primero, al hecho de la poca edad de la mayoría de los informantes y que, tal vez por eso, no hayan tenido la oportunidad de vivir contextos semejantes a lo ejemplificado en la oración del par *banco de arena*; segundo, porque el par de oraciones uno fue el primero analizado por los informantes y, por ello, puede ser que por la inseguridad al contestar a las preguntas no hayan reflexionado con mucha atención sobre las dos unidades en los respectivos contextos presentados (Pereira, 2018; Rodrigues-Pereira; Nadin, 2020).

el 94,4 % de respuestas negativas, lo que nos garantiza considerarla un caso de homonimia. Por un lado, porque estamos considerando, en una perspectiva sincrónica, el criterio de la divergencia semántica entre las UL en términos de significado para los informantes. Por otro, porque podemos definirla como un caso de homonimia porque son unidades léxicas de categorías gramaticales diferentes: sustantivo y verbo, respectivamente, lo que Biderman (1978) define como *homónimos sintácticos* y Zavaglia (2003) clasifica como *homonimia categorial*.

Para el grupo de la UL *carta*, con los dos pares de oraciones, identificamos un caso de polisemia y otro de homonimia. En i) “¿No has leído la **CARTA** que te envié?” (DEECMG, 2005); “En la **CARTA** pone el precio de cada plato” (DEECMG, 2005), tuvimos el 64,8 % de respuestas positivas, datos que confirman la relación de sentido para los informantes y, por ello, un caso de polisemia. Ya en ii) “Con la ayuda de la **CARTA** náutica podemos viajar tranquilos” (adaptado de DEECMG, 2005); “La baraja española tiene cuarenta **CARTAS**” (DEECMG, 2005), el 68,5 % de respuestas fue negativo, lo que confirma el valor homonímico en este caso.

La UL *gato*, con dos pares de oraciones y el 83,3 % y el 69,4 % de respuestas negativas, quedó claramente confirmada como caso de homonimia, pues tanto el gato/animal como el gato/aparato que sirve para levantar el auto y el gato/persona nacida en Madrid, para la mayoría de los informantes, no hay relación de sentidos.

Para *taco*, con el 88 % de las respuestas positivas para el par i): “Mi amigo mexicano prepara unos **TACOS** de pollo muy ricos” (DEECMG, 2005), *tortilla de maíz*; “Como no tengo tiempo de ir a comer, me echaré un **TACO** por acá” (DEECMG, 2005), *comida ligera*, tenemos un caso típico de polisemia, pues el alto grado de relación de

sentido para los informantes nos posibilita tal afirmación. Ya para el par ii): “Una hoja del **TACO** del calendario” (DEECMG, 2005); “Como aperitivo tomamos unos **TACOS** de jamón” (DEECMG, 2005), con el 93,5 % de respuestas negativas, verificamos un caso de homonimia; confirmando, pues, la falta de relación de sentido para la mayoría de los informantes.

En el grupo *coma*, en los dos pares, obtuvimos el 68,5 % y el 98,1 % de respuestas negativas, ratificando el fenómeno homonímico para esta UL de la lengua española. Tanto en el par i) *coma/estado patológico*; *coma/signo ortográfico*, como en el par ii) *coma/estado patológico/*; *coma/verbo*, para la mayoría de los informantes, no hay relación de sentido entre las UL.

Para el grupo *llama*, solamente el par i) “Las **LLAMAS** del incendio arrasaron varias hectáreas” (DEECMG, 2005); y “Manténían viva las **LLAMAS** de su amor” (DEECMG, 2005), tenemos un caso de polisemia, pues la mayoría de los informantes, el 88,9 %, están de acuerdo en que hay relación de sentido entre las lexías. Los pares ii) y iii), con el 86,1 % y el 97,4 %, respectivamente, se clasifican como casos de homonimia por la no relación de sentido existentes.

En *terrazza*, con solo un par: “Mi casa tiene **TERRAZA**” (DEECMG, 2005); y “Nos sentamos en una mesa de la **TERRAZA** del bar” (DEECMG, 2005), obtuvimos el 99,1 % de respuestas positivas, lo que se clasifica, también, como un caso típico de polisemia, por la relación de sentido que hay entre los dos usos.

Con el grupo *canto*, a partir de los tres pares de oraciones utilizadas en esta investigación: i) “Creo que está dotada para el **CANTO**” (DEECMG, 2005); y “Estudio **CANTO** en el conservatorio” (DEECMG, 2005); ii) “Estudio **CANTO** en el conservatorio” (DEECMG, 2005); y “Cuidado, no te des con el **CANTO** de la mesa” (DSEE,

2006); **iii**) “Estudio *CANTO* en el conservatorio” (DEECMG, 2005); y “Me siento feliz cuando *CANTO* a mis alumnos”, identificamos importantes datos que carecen de reflexiones más profundas, sobre todo porque los informantes se posicionan en una dirección contraria a algunos criterios para el establecimiento de la homonimia, como los que ya hemos demostrado en Pereira (2018) y Rodrigues-Pereira; Nadin (2020).

Por los datos obtenidos, en i) detectamos que el 98,1 % de las respuestas son positivas y solo el 1,9 % negativas, lo que nos permite comprender *canto* en ese contexto como un caso típico de polisemia. Para el par ii), con el 92,6 % de respuestas negativas, también tenemos otro caso típico, pero, ahora, de homonimia. Ya para el par iii), aun para el criterio formal de la homonimia sintáctica (Biderman, 1978), así como el de la homonimia categorial explicada por Zavaglia (2003), sean una realidad importante de ser considerada, dependiendo de los objetivos de la investigación, los datos evidencian la estrecha relación de sentido existentes entre *canto/música/sustantivo* y *canto/cantar/verbo*, puesto que ambas formas, aunque pertenecientes a categorías gramaticales diferentes, mantienen relación de sentido para la mayoría de los informantes.

Los datos de esta investigación, así como los de Pereira (2018) y Rodrigues-Pereira; Nadin (2020) nos han posibilitado discutir con respecto al valor de contenido que una UL posee frente a su forma ortográfica y categorial. En realidad, desde 2017 cuando empezamos a inventariar los candidatos a homónimos para el *DMHE*, y para ellos tuvimos que revisar numerosos criterios para el establecimiento de homonimia para la *LEXPED*, hemos reflexionado teórica y metodológicamente¹⁸, en ese contexto, a partir de la hipótesis de que el

¹⁸ Germain (1981), Guiraud (1989), Polguère (2018), Silva (2010), Trujillo (1988), Ullmann (1964), entre otros.

sentido precede a la forma. Con datos semejantes a los que presentamos con esta investigación semántica, hemos identificado usos que, de forma pertinente, confirman nuestra afirmación. Sin embargo, queda aún abierta la cuestión para futuros trabajos que puedan sumarse a los ya existentes sobre el tema.

Por último, el grupo *tienda* con sus dos pares de oraciones: i) “En esta calle hay muchas **TIENDAS** de zapatos” (adaptado de DSEE, 2006); y “Los excursionistas llevaban varias **TIENDAS** para acampar en la montaña” (DSEE, 2006); ii) “En esta calle hay muchas **TIENDAS** de zapatos” (adaptado de DSEE, 2006); y “Señor Juan, por favor, **TIENDA** la cama”; constatamos, también, dos casos de homonimia, pues la mayoría de los informantes contestaron la encuesta con respuestas negativas, como se ve por los datos porcentuales: el 61,1 % y el 97,2 %, respectivamente.

Como se percibe por lo expuesto, de los veintiún (21) pares de oraciones, quince (15) fueron identificados como casos de homonimia, es decir, un porcentaje de 71,5 % de los datos; para la polisemia, tuvimos un total de seis (06), con el 28,5 %. Así como Silva (1986, *apud* Biderman, 1991, p. 288), identificamos, de igual modo, un alto grado de acuerdo entre los hablantes en la discriminación de estas dos categorías, pues tuvimos una media de un 80,5 %, si sumamos todos los porcentajes negativos de la Tabla I y dividimos por el total de quince (15) pares, cuyas UL fueron identificados como ULH. De la misma forma que el investigador antes citado, no tuvimos gran desacuerdo entre los hablantes, de forma que nunca hubo un acuerdo inferior a 60 %, como se ve por los datos demostrados en esta investigación.

CONCLUSIONES

Considerando los objetivos de este texto, así como los planteados para la investigación semántica que nos ha dado la oportunidad de reflexionar de forma teórica y metodológica al respecto de algunos hechos homonímicos, hemos podido alcanzar positivamente nuestras intenciones investigativas para el establecimiento de la homonimia destinada a diccionarios pedagógicos a partir de criterios de naturaleza sincrónica, como el que demostramos en este texto.

En efecto, lo que se confirma es que si partimos del principio de la divergencia semántica (Ullmann, 1964), del criterio de la consciencia lingüística del hablante (Werner, 1982), así como de las constataciones de Silva (1986, *apud* Biderman, 1991), de Pereira (2018), Rodrigues-Pereira; Nadin (2020) y de los datos de la investigación semántica que presentamos, podemos perfectamente considerar como casos de homonimia, para la LEXPED, todas las UL que, en una perspectiva sincrónica, no poseen relación de sentido.

Los datos aquí expuestos corroboran la aseveración ya discutida anteriormente de que, para los estudiantes de una o más lenguas, en especial los de lengua extranjera, lo que importan son los sentidos que una UL posee en un determinado estado de la lengua, para que puedan expresar lo que desean de forma competente en la lengua en cuestión.

REFERENCIAS

- ALOCHIS, Ivana. **Diccionario temático de adivinanzas – volumen II:** utensilios, construcciones y manufacturas. 1. ed. Córdoba: El Rescoldo, 2007.
- BIDERMAN, Maria Teresa Camargo. A ciência da Lexicografia. **Revista Alfa**, São Paulo, v. 28 (Suplemento), p. 1-26, 1984.

BIDERMAN, Maria Teresa Camargo. **Dicionário Didático de Português**. 2. ed. São Paulo: Ática, 1998.

BIDERMAN, Maria Teresa Camargo. Polissemia versus homonímia. *In: SEMINÁRIO DO GRUPO DE ESTUDOS LINGÜÍSTICOS DO ESTADO DE SÃO PAULO*, 38., 1991, Franca. **Anais [...]**. Franca: Unifran – União das Faculdades Franciscanas, 1991. p. 283-290. Disponível em: http://www.gel.hospedagemdesites.ws/arquivo/anais/1306936856_43.biderman_ma.tereza.pdf. Acesso em: 23 ago. 2023.

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

BIDERMAN, Maria Teresa Camargo. Unidades complexas do léxico. *In: RÍO-TORTO, G. et al. (org.) Estudos em homenagem ao Professor Doutor Mário Vilela*. Porto: Faculdade de Letras da Universidade do Porto, 2005. p. 747-757.

GERMAIN, Claude. **La semântica funcional**. Madrid: Editorial Gredos, 1981.

GONZÁLEZ, Maldonado Concepción. **DEECMG - Diccionario de español para extranjeros** – Con el español que se habla hoy en España y en América Latina. Coordinación y proyecto editorial Concepción Maldonado. São Paulo: edições SM, 2005.

GUIRAUD, Pierre. **A Semântica**. 5. ed. Rio de Janeiro: Editora Bertrand Brasil S. A., 1989.

GUTIÉRREZ CUADRADO, Juan; PASCUAL RODRÍGUES, José Antonio. **DSEE - Diccionario salamanca – español para extranjeros**. Madrid: Santillana Educación, 2006.

LYONS, John. **Semântica**. v. 1. Lisboa: Editorial Presença/Martins Fontes, 1977.

MORANTE VALLEJO, Roser. **El desarrollo del conocimiento léxico en segundas lenguas**. Madrid: Arco Libros, S.L., 2005.

PEREIRA, Renato Rodrigues. **O dicionário pedagógico e a homonímia: em busca de parâmetros didáticos**. 2018, 209f. Tese (Doutorado em Linguística e Língua Portuguesa) – Universidade Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências e Letras, Araraquara, 2018.

POLGUÈRE, Alain. **Lexicologia e semântica lexical**: noções fundamentais. Tradução de Sabrina Pereira de Abreu. São Paulo: Contexto, 2018.

PORTO DAPENA, José Álvaro. **Manual de técnica lexicográfica**. Madrid: ARCO/LIBROS, S. A, 2002.

RODRIGUES-PEREIRA, Renato; NADIN, Odair. Luiz. Bases teóricas e metodológicas para o inventário de candidatos a homônimos destinados a dicionários pedagógicos. **Signótica**, Goiânia, v. 32, 2020. Disponível em: <https://revistas.ufg.br/sig/article/view/65314/36156>. Acesso em: 24 ago. 2023.

RODRIGUES-PEREIRA, Renato; ZACARIAS, Regiani Aparecida Santos; NADIN, Odair Luiz. Léxico, ensino e suas interfaces. **Revista GTLex**, Uberlândia, v. 5, n. 1, p. 6-22, 2020. Disponível em: <https://doi.org/10.14393/Lex9-v5n1a2019-1>. Acesso em: 24 ago. 2023.

SILVA, Augusto Soares da Silva. Palavras, significados e conceitos: o significado lexical na mente, na cultura e na sociedade. **Cadernos de Letras da UFF – Dossiê: Letras e cognição**, n. 41, p. 27-53, 2010. Disponível em: https://www.academia.edu/34275927/Palavras_significados_e_conceitos_o_significado_lexical_na_mente_na_cultura_e_na_sociedade. Acesso em: 8 out. 2021.

TRUJILLO, Ramón. **Introducción a la semántica española**. Madrid: ARCO/LIBROS, S. A., 1988.

ULLMANN, Stephen. **Semântica**: uma introdução à ciência do significado. Tradução de J.A. Osório Mateus. Lisboa: Fundação Calouste Gulbenkian, 1964.

WERNER, Reinhold. Léxico y teoría general del lenguaje. In: HAENSCH, G. *et al.* **La Lexicografía**. De la Lingüística teórica a la Lexicografía práctica. Madrid: Editorial Gredos, 1982. p. 20-94.

ZAVAGLIA, Cláudia. Ambiguidade gerada pela homonímia: revisitação teórica, linhas limítrofes com a polissemia e proposta de critérios distintivos. **DELTA: Documentação de Estudos em Linguística Teórica e Aplicada**, São Paulo, v. 1, n. 19, p. 237-266, 2003.

ZAVAGLIA, Cláudia. **Análise da Homonímia no português**: tratamento semântico com vistas a procedimentos computacionais. 2002. 574 f. Tese (Doutorado em Linguística e Língua Portuguesa) – Universidade Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências e Letras:Faculdade de Ciências de Letras - UNESP, Araraquara, 2002.

NÓBREGA [AOS MORADORES DE PERNAMBUCO], BAÍA [5 DE JUNHO] 1552 [A00_0007 p. 334].

MEL DE PAU

Mel das abelhas uruçú, jataí e outras que o ajuntam em ocós das árvores. [...] entãõ fe lhe lãcem duas oytavas de fenne, com que dê huma leve fervura, e tirado do fogo fe coe, e torne a elle a purificar com fete, ou oyto onças de mel de páo, dando algumas fervuras, e fendo neccffãrio coado, fe coe, e guarde para o ufo. LUIS GOMES FERREIRA (1735) [1735], DAS OBSTRUCÇÕES [B00_0030 p. 82].

MEL DE ENGENHO

Caldo da cana cozido e grosso que se apura para ir às formas a purgar. MEL DE ENGENHO Ouve-se falar n'elle tão a miudo que entender-se-á talvez, que esta é uma das grandes manufacturas. Examinando o caso está visto, que as maiores quantidades não passam de algumas frasqueiras d'elle, que os fabricantes vendem á razão de 320 rs. cada frasco para supprir a falta do assucar. ALEXANDRE RODRIGUES FERREIRA [séc. XVIII], 2.^a PARTE: BAIXO RIO

PARTE III TERMINOLOGIA

A LINGUAGEM DO PATRIMÔNIO CULTURAL DE BENS MÓVEIS EM PAPEL: CONSTRUÇÃO DE UMA BASE DE DADOS TERMINOLÓGICA

*Anna Maria Becker Maciel
Cleci Regina Bevilacqua*

INTRODUÇÃO

O crescente interesse pela conservação-restauração do patrimônio cultural da humanidade revelado na internet pela proliferação de oficinas, centros de pesquisa, cursos técnicos e acadêmicos, associações nacionais e internacionais parece ser uma das características da nossa época. Também no Brasil esse interesse se faz sentir; conservadores-restauradores buscam qualificação e exigem leis que reconheçam sua profissão (Regulamentação..., 2020), universidades criam programas de graduação e pós-graduação e grupos de pesquisa investigam a organização e os fundamentos teórico-metodológicos da área (Redini, 2019).

A efervescência desse domínio de saber e prática até há pouco confiado a amadores tende a se consolidar como ciência e faz surgir um conjunto grande de textos de todos os tipos e formatos, temas e áreas relacionadas com a preservação dos mais variados bens do patrimônio cultural. Um imenso cabedal de informações multilíngues

atrai a atenção de especialistas e leigos. Diante da carência de glossários e dicionários especializados para auxiliar a compreensão de sua leitura, o Grupo Terminológico Cone Sul (TERMISUL), do Instituto de Letras, UFRGS, se propôs pesquisar e divulgar o *modus-dicendi* da área. Por ora, limitamos o escopo de nosso trabalho à terminologia da Conservação-Restauração de bens culturais móveis em suporte papel.

De 2016 a 2021, o TERMISUL desenvolveu dois projetos de pesquisa relacionados ao tema. O primeiro intitulou-se *A linguagem do patrimônio cultural brasileiro: conservação dos bens culturais móveis* (2016-2019) e teve o objetivo de coletar os termos da área. O segundo denominou-se *Conservação dos bens culturais móveis em papel: identificação e representação de unidades fraseológicas especializadas (UFES)*, realizado entre 2019 e 2021. A aplicação dos resultados de ambos foi a construção de uma base de dados terminológica (BDT PAPEL) *on-line* multilíngue para registro e representação de termos e Unidades Fraseológicas Especializadas Eventivas (UFEEs)¹, tendo como público preferencial aprendizes e profissionais da área de Conservação e Restauração bem como tradutores, redatores, revisores de textos, aprendizes e profissionais da área de Conservação e Restauração e demais interessados.

As justificativas para levar a cabo tais projetos residem no fato de que, na esfera acadêmica, a conservação, a preservação e a restauração do patrimônio cultural, praticadas tanto por amadores como por especialistas, começaram a se configurar como área especializada fundamentada em quadro teórico a partir das últimas décadas com a criação de cursos universitários. Seus especialistas utilizam uma linguagem multifacetada que recorre a terminologias das Ciências Humanas, das Ciências Exatas e de técnicas milenares. Nesse amálgama de saberes e competências, a comunicação, quer entre pares, quer com

¹ Disponível em: <https://www.ufrgs.br/termisul/papel/>.

o público interessado, gera uma linguagem com maneiras de dizer e terminologias oriundas de diversos domínios, culturas, tradições, línguas e mídias, causando confusão e falhas de compreensão.

Tal diversidade dificulta a comunicação até mesmo entre os falantes de uma mesma língua. A demanda de produtos terminológicos registrada no documento de Pavia (ICOMOS, 1997) continua, e a necessidade da unificação da terminologia segue discutida em seminários internacionais.

No Brasil, essa lacuna também motivou pesquisas do léxico da área principalmente a partir da criação de cursos sobre a conservação do patrimônio cultural nas instituições de ensino superior como na Universidade Federal de Minas Gerais (UFMG), na Universidade Federal do Rio de Janeiro (UFRJ) e na Universidade Federal de Pelotas (UFPel), entre outras.

Acresce que tanto internacional como nacionalmente, nos projetos orientados para a terminologia, a contribuição do linguista e/ou terminólogo raramente é mencionada. São os especialistas em conservação e restauração que se dedicam a inventariar os termos que lhes parecem próprios desse campo de saber e atividade. Conheçamos poucos trabalhos que registram a contribuição do terminólogo como o de Xarrié (2015) que descreve a linguagem da conservação artística, do ponto de vista multidisciplinar, acadêmico e internacional. No Brasil, tanto quanto sabemos, temos os seguintes trabalhos: *Terminologia em conservação de bens culturais em papel* (Bojanoski, 2018) e *Dicionário ilustrado da conservação de documentos gráficos* (Barbosa, 2017). Salientamos, portanto, a pertinência da realização do projeto que ora relatamos desenvolvido na área da terminologia contemporânea e que associa linguistas do Grupo Termisul do Instituto de Letras, UFRGS, e especialistas do Curso de Conservação e Restauração de Bens Culturais Móveis, UFPel.

Neste texto apresentamos as bases teóricas que sustentam ambos os projetos, as etapas metodológicas da pesquisa e do tratamento dos dados e os resultados alcançados que culminaram na construção da base de dados. Esperamos, assim, poder enfatizar algumas questões teórico-metodológicas implicadas na elaboração de um produto terminográfico multilíngue e *on-line* contendo unidades especializadas de uma área em consolidação e que ainda requer reconhecimento em nosso país, a Conservação e Restauração de Bens Culturais.

1. A LINGUAGEM DA CONSERVAÇÃO DOS BENS CULTURAIS MÓVEIS

Nossa pesquisa abrange a linguagem da conservação dos bens culturais móveis a partir de documentos gráficos manuscritos e/ou impressos em suporte papel como folhas avulsas, livros, partituras, cartazes, folhetos, revistas e jornais. Desenhos, mapas, plantas arquitetônicas, fotografias, selos, cédulas e obras tridimensionais que utilizam o mesmo suporte ainda que consideradas bens culturais não foram contemplados.

O conceito de bem cultural é muito discutido, diante da multiplicidade de definições encontradas, optamos pela definição do Código de Ética do Conservador-Restaurador (Código..., 1996, p. 1): “Por bens culturais entendemos aqueles objetos a que a sociedade atribui particular valor artístico, histórico, documental, estético, científico, espiritual ou religioso”. Todavia, ponderamos que é preciso,

[...] ter sempre em vista que se trata de uma concepção em processo, e que envolve perspectiva multidisciplinar, considerando que cada período da história está voltado para determinados interesses que vão, de alguma forma, alterar e interferir no

significado que podemos dar ao termo bem cultural (Guedes; Maio, 2016).

Ainda de acordo com o mesmo Código de Ética acima citado, entendemos conservação-restauração como “o conjunto de práticas específicas, destinadas a estabilizar o bem cultural sob a forma física em que se encontra, ou, no máximo, recuperando os elementos que o tornem compreensível e utilizável, caso tenha deixado de sê-lo” (Código..., 1996, p. 1).

Quanto ao objeto linguístico, buscamos palavras que revelam um significado específico na conservação-restauração dos bens culturais móveis em suporte papel. Limitamos nossa coleta às palavras consideradas termos, isto é, Unidades Terminológicas (UTs), e UFEEs. As UTs configuram conceitos da área especializada. Conforme Cabré (1999, 2001), são entendidas como “unidades léxicas, que se descrevem como unidades denominativo-conceituais, dotadas de capacidade de referência, que podem exercer diferentes funções (referencial, expressiva, conativa) e que, integradas no discurso, constituem a quase totalidade dos argumentos das orações” (Cabré, 1999, p. 132, tradução nossa). Por exemplo, *velatura* definido como “ato de colar sobre o verso de documentos executados sobre suporte de papel uma folha de papel japonês com o uso de cola metilcelulose, tendo o objetivo de conferir-lhe maior resistência física” (Spinelli Júnior, 1997, p. 77). Como exemplos de termos simples da área temos *acervo*, *arquivo*, *espátula*, *gramatura* e de termos compostos, *água desmineralizada*, *cola metilcelulose*, *máquina obturadora de papel*.

Por sua vez, as UFEEs são definidas como unidades pluriverbais, formadas por um Núcleo Eventivo (NE) de caráter verbal ou nominalizado e um Núcleo Terminológico (NT) e referem-se a processos e ações da área (Bevilacqua, 2004). Segundo a autora, en-

tre esses dois núcleos se estabelecem relações de tipo sintático, mas principalmente de caráter semântico, determinadas pelas condições pragmático-discursivas, o que confere à combinação um caráter estável, i.e., de unidades semifixas. Nesse sentido, as UFEEs são unidades que se conformam pelo e no discurso em que ocorrem, passando a ter valor fraseológico e especializado pelas características do texto em que são utilizadas, principalmente por aspectos pragmáticos como a temática e a situação comunicativa (interlocutores envolvidos, grau de especialização e finalidade dos textos). Nessa perspectiva, a UFEE adquire valor especializado e sua estrutura subjacente pode realizar-se superficialmente em três estruturas, quais sejam formadas por verbo (*preservar acervo*), nominalização (*preservação de acervo*) e participípio (*acervo preservado*). Nos limites do projeto aqui referido foram identificadas as unidades de estrutura verbal (*abrigar coleção, restaurar documento, remover sujidade*) e nominal (*remoção de adesivo, oxidação de celulose, remoção de sujidade*).

Feitos os esclarecimentos sobre algumas características da linguagem da área em estudo e das unidades que constituíram o objeto de investigação dos projetos relatados – UTs e UFEEs – apresentamos, na sequência, a fundamentação teórica que sustentou as pesquisas.

2. PROPOSTA TEÓRICA

Para sustentar a pesquisa da linguagem de uma comunidade profissional de especialistas de formação acadêmica e técnica variada, cujo desempenho profissional se alimenta em fontes de diferentes origens e idiomas, é preciso encontrar um arcabouço teórico-metodológico que, na multiplicidade das áreas temáticas e das línguas envolvidas, consiga dar conta do aspecto social, comunicativo, linguístico e cognitivo da terminologia utilizada. Para tanto, em um projeto que

contempla língua materna e línguas estrangeiras, recorreremos às teorias que integram a sintaxe, a semântica e a pragmática também sob o prisma da tradução e que consideram a linguagem especializada uma expressão da língua comum em um contexto ativado por especificidades próprias de dada ciência, técnica ou habilidade. Nesse quadro, destacamos os pressupostos compartilhados pela Socioterminologia (ST), pela Teoria Comunicativa da Terminologia (TCT), pela Teoria Sociocognitiva da Terminologia (TST), pela Terminologia Linguístico-Textual (TLT), pelos estudos da Fraseologia Especializada e pelos Estudos de Tradução atuais e pela Linguística de *Corpus*.

A ST (Gaudin, 1993; 2005) chama atenção para a influência do contexto sociocultural da comunicação, sublinhando que o estatuto social e profissional dos interlocutores se reflete na configuração da linguagem especializada e determina a multiplicação de conceitos e consequente variação da terminologia usada. Assim, a ST defende a análise das condições de circulação do termo no seu real funcionamento (Fausltich, 2006).

Paralelamente, a TCT (Cabré, 1999; 2001) prioriza a multidimensionalidade da unidade de conhecimento especializado e salienta o componente linguístico, cognitivo e sociocomunicativo do termo visto como palavra da língua comum cuja especificidade é ativada por condicionamentos da área. Os principais postulados da TCT podem ser resumidos na visão do caráter lexical e comunicativo da unidade terminológica transmissora de conhecimento especializado, cuja poliedricidade e variação resultam do contexto situacional de uso em determinada área temática.

Ao sublinhar o dinamismo da comunicação especializada, a TST (Temmerman, 2000) explica que a especificidade da linguagem advém de uma série de condicionamentos que, ao correr do tempo,

se refletem nas características cognitivas e pragmáticas que estruturaram os modelos cognitivos selecionados pelo especialista para compreender a realidade da área temática. O postulado máximo da TST é o entendimento do conceito como uma unidade de compreensão de estrutura prototípica resultante do mecanismo da pragmática do discurso especializado.

Recorremos também a Krieger (2004; 2008) que enfatiza a relevância da Terminologia Linguístico Textual (TLT) para a descrição da organização narrativa dos textos especializados e seu patamar discursivo, o que permite entender “[...] o funcionamento da linguagem em distintos contextos de uso para identificar o comportamento dos termos, suas variantes, sinónimas” (Krieger, 2017, p. 3).

Quanto à tradução, seguimos a perspectiva comunicativa de Hurtado Albir (2001). Segundo essa autora, a tradução é definida como “processo interpretativo e comunicativo que consiste na reformulação de um texto com os meios de outra língua que se desenvolve em um contexto social e com uma finalidade determinada” (Hurtado Albir, 2001, p. 41, tradução nossa). Nessa perspectiva, considera-se a tradução como ato de comunicação, operação textual e atividade cognitiva. Assim, no processo tradutório, é preciso analisar os interlocutores implicados na tradução (quem e para quem), as circunstâncias e fatores históricos tanto da língua de partida como da língua de chegada (quando e onde), além dos aspectos linguísticos (como).

Também adotamos a definição de equivalência proposta pela autora que apresenta uma definição de equivalência flexível e dinâmica, considerando-a como um vínculo entre os textos originais e suas traduções, embora esse vínculo também possa ocorrer entre unidades dos textos (unidades de tradução), maiores ou menores (Hurtado Albir, 2001, p. 223). Desse modo, para a identificação de equivalência,

é necessário levar em conta os seguintes aspectos: o contexto sócio-histórico do texto de chegada, o contexto linguístico do texto de chegada, a função ou finalidade do texto traduzido, o gênero textual e a modalidade da tradução (escrita, audiovisual, etc.). Tais parâmetros foram aplicados para a identificação dos equivalentes dos termos e UFEs, considerando que nos projetos aqui relatados se trata de busca de equivalentes para um produto terminográfico e não para a tradução de um texto; são equivalentes que poderão ser utilizados em uma tradução e que, portanto, servem de base para as pesquisas dos tradutores no momento da realização de uma tradução.

A esse amplo quadro referencial, associamos a abordagem teórico-metodológica preconizada pelos autores Biber, Conrad e Reppen (1998), Stubbs (1996; 2001) e Berber Sardinha (2004) que postulam a configuração do significado no contexto real da comunicação e o desenvolvimento da pesquisa linguística com *corpus*.

3. ETAPAS METODOLÓGICAS

Sintetizamos, a seguir, as etapas metodológicas que permitiram o desenvolvimento da pesquisa linguística e a criação da BDT PAPEL a partir do português brasileiro com equivalentes em espanhol, francês, inglês, italiano e russo, disponibilizada na página do TERMISUL. Destacamos a construção dos *corpora* de pesquisa, os critérios para a identificação e coleta das entradas, a estruturação e a navegabilidade do sistema de apresentação e de consulta.

A primeira etapa foi a construção de um *corpus* representativo de textos da área em foco em língua portuguesa. Os textos foram coletados de *sites* reconhecidamente autorizados e indicados pela pesquisadora-colaboradora da UFPel, Profa. Dra. Silvana Bojanoski, como a

Associação Brasileira de Conservadores e Restauradores (ABRACOR), universidades, museus, periódicos científicos e eventos acadêmicos de áreas afins ao tema da conservação do papel. Foram coletados relatórios de pesquisa, dissertações, teses, normas técnicas, manuais e artigos científicos de especialistas dirigidos a seus pares, publicados em *sites* institucionais. Para a construção dos *corpora* nas línguas estrangeiras foram seguidos critérios semelhantes a fim de obter *corpora* de contraste. Contudo, não conseguimos uma distribuição paralela dos textos; coletamos o que foi possível na época e acreditamos em sua representatividade, pois, além da diversidade de gêneros, chegamos a cerca de 1 milhão de palavras para cada um. O número de textos e os países incluídos para cada *corpus* é apresentado no quadro 1.

Quadro 1 – Número de textos, de palavras e países de cada *corpus*

Língua	Textos	Países
Português (ptPP)	161	Brasil
Espanhol (arPP, chPP, esPP, mxPP e uyPP)	125	Argentina, Chile, Espanha, México e Uruguai
Francês (bePP, caPP, frPP, swPP)	127	Bélgica, Canadá, França e Suíça
Inglês enPP	161	Textos originais em inglês
Italiano (itPP)	75	Itália
Russo (ruPP)	64	Rússia e Bielorrússia

Fonte: as autoras

A etapa seguinte foi a extração dos termos para a qual foi utilizado, em um primeiro momento, o programa Sketch Engine (Killgarriff *et al.*, 2014) e o recurso *Keywords/Terms*, disponível em *Keywords* (palavras-chave), que comparou a frequência das palavras do *corpus* de estudo (ptPP) com a frequência dessas palavras no *corpus* de língua geral de referência (Portuguese Freeling 1.3.2), disponibilizado no próprio programa. A partir dos resultados, foram aplicados

critérios de corte de frequência (igual ou superior a 10) e a exclusão de nomes próprios (pessoas, organizações e lugares), palavras estrangeiras, abreviaturas e símbolos químicos e outros tipos de ruídos.

A seguir organizaram-se os dados nas cinco categorias propostas na árvore domínio da área da conservação de documentos gráficos em formato papel (Bojanoski, 2018). Tal categorização permitiu perceber a multidisciplinariedade da linguagem utilizada tanto quanto aos conceitos fundamentais (*conservação, restauração, recuperação, preservação*) como quanto à operacionalização dos procedimentos (*salvaguarda, manuseio, limpeza, encolagem*), uso de produtos, instrumentos e equipamentos (*mesa de sucção, óculos de proteção*).

Para a identificação dos contextos de uso, foi utilizado o programa AntConc (Antony, 2018), principalmente o recurso *Concordance*. Como resultado final, foram selecionados cerca de 300 termos simples e sintagmáticos, dos quais trazemos exemplos que enfatizam a multidimensionalidade das disciplinas implicadas e que se estendem desde as áreas exatas até as humanas, passando por tecnologias e práticas tradicionais: *acervo, acidez, ácido sulfúrico, foxing, máquina obturadora de papel (MOP), tinta ácida, karibari*. Além disso, a lista foi revisada por Silvana Bojanoski, especialista na área, que sugeriu a inserção de alguns termos importantes na área para complementar a lista anterior.

Para a seleção das UFEEs em português, realizada no âmbito da pesquisa de mestrado de Arcos (2019), foi utilizado o mesmo *corpus* em português e a ferramenta *Word Sketch* também do *Sketch Engine*. Assim, levantaram-se os padrões colocacionais para os termos mais salientes do *corpus*. Por exemplo, para o termo *papel*, foram indicadas formas nominalizadas como *acidificação de, envelhecimento de* e verbos como *utilizar*.

Dos resultados gerados, foram selecionadas as unidades de frequência absoluta igual ou superior a 10. A análise de seus contextos

permitiu verificar se as estruturas resultantes se configuravam como UFEEs da área, isto é, se faziam referência aos processos e ações específicos da área. Como resultado final, foram identificadas 160 UFEEs, entre as quais estão: *deterioração de documento, higienização de livro, limpeza de acervo, salvaguarda de patrimônio documental e tratamento do papel.*

Para a busca dos equivalentes dos termos, partiu-se da proposta de equivalência de Hurtado Albir (2001, p. 223), conforme referido anteriormente. Os equivalentes foram identificados nos *corpora* das línguas contempladas a partir da geração de uma lista inicial de palavras-chave (*Keywords/Terms* do *Sketch Engine*), utilizando-se *corpora* de referência para cada uma das línguas e aplicando-se corte de frequência igual ou maior que 10. Após a análise e cotejamento de contextos dos termos em português e nas línguas estrangeiras, foi feita sua validação. Essa busca foi complementada com a análise dos contextos dos termos e UFEEs com o uso da ferramenta *Concordance* do *AntConc*. Nos casos de ausência de equivalentes nos *corpora*, optamos por buscas em *sites* externos, observando critérios para assegurar a confiabilidade dos dados, tais como a idoneidade das instituições, publicações e eventos.

4. ESTRUTURA DA BDT PAPEL

A estrutura da BDT PAPEL seguiu as linhas gerais do padrão tradicionalmente adotado pelo TERMISUL em suas bases terminológicas (Maciel, 2013), tanto no que se refere à organização, manutenção e consulta como à definição da ficha terminológica geradora da entrada. Assim, a base compõe-se de dois módulos: a) módulo trabalho de acesso restrito à administração e b) módulo de consulta de livre acesso (figura 1). Trata-se de um sistema amigável que não exige

nenhum conhecimento linguístico ou informático do consulente. A busca pode ser efetuada por palavra inteira ou truncada e também por ordem alfabética a partir de qualquer uma das línguas contempladas. Foram inseridas 529 entradas em português, 434 em inglês, 511 em espanhol, 533 em francês, 513 em italiano e 466 em russo, perfazendo o total de 2.986 entradas. A manutenção e atualização das informações são feitas facilmente pelos administradores da base.

Figura 1 – Módulos da BDT PAPEL

Para conhecer a documentação de BDT PAPEL clique [aqui](#).

Para fazer uma busca por palavra [aqui](#).

Para fazer uma busca alfabética [aqui](#).

Para editar a base de dados [aqui](#).

Módulo de consulta

Módulo de trabalho

Fonte: TERMISUL

Cada entrada consta de ficha terminológica com sete campos. O primeiro, *entrada*, é reservado à UT ou UFEE, seguido pela língua da entrada; o terceiro é seu contexto e código da fonte referencial no *corpus* de pesquisa a ser encontrado nos catálogos disponibilizados na documentação da BDT. O quarto campo, *outras formas*, indica algum tipo de variação morfológica das unidades léxicas ou à explicitação de siglas (ex: *EPI e equipamento de proteção individual*) juntamente com seu contexto. O campo *ver também* registra as relações de sinonímia e hiponímia da entrada (tipos de); são *hiperlinks* que remetem às fichas terminológicas completas (ex.: *papel japonês* remete a *papel washi*, *limpeza de acervo* remete para *higienização de acervo*); enquanto *notas* é o campo reservado para a explicação de siglas, abreviaturas do contexto ou outra informação relevante sobre a entrada. O campo *equivalentes*

contém as remissivas do português para os equivalentes de cada uma das línguas estrangeiras e remissivas da língua estrangeira para o português. Finalmente, *comentários* é o espaço para interlocução entre os pesquisadores; campo presente apenas na base de trabalho. A figura 2 ilustra a entrada do termo *alvura*.

Figura 2 – Entrada do termo *alvura*

TERMO: ALVURA

Língua: Português

Contexto:

O branqueamento é um tratamento químico, em vários estágios, que se dá à celulose após o cozimento, a depuração e a lavagem, à pasta mecânica depois do desfibramento [...] com a finalidade de descolorir ou remover os materiais corantes não-celulósicos existentes na massa, aumentando a **alvura** do produto final [...] (ptPP151)

Ver também: ➡ Relação de sinonímia e/ou tipos de

brancura

Equivalente(s) em Inglês :

whiteness

Equivalente(s) em Espanhol :

blancura

Equivalente(s) em Francês :

blancheur

Equivalente(s) em Italiano :

bianchezza

Equivalente(s) em Russo :

белизна [belizna]

Link para os equivalentes

[Edite este Termo](#)
 [Veja outro Termo](#)
 [Busca dos Termos](#)

[Anterior](#)
 [Módulo de trabalho](#)
 [Sair](#)

Fonte: TERMISUL

5. RESULTADOS E ASPECTOS RELEVANTES

Lançando um olhar retrospectivo na caminhada para a consecução dos dois projetos, destacamos a relevância do trabalho mul-

tilíngue e da pesquisa na perspectiva da Terminologia Linguístico-Textual, tradicionalmente presentes na reflexão teórica e sua aplicação no grupo TERMISUL desde sua fundação em 1991. Trabalhar no ambiente multilíngue supõe triplicar, quadruplicar ou quintuplicar o trabalho, dependendo do número de línguas incluídas. É preciso lidar com diferenças culturais e linguísticas que se refletem tanto na constituição dos termos como das fraseologias especializadas e no seu uso. Por outro lado, a confluência das diferentes línguas na configuração da linguagem especializada permite investigar a multidimensionalidade das unidades léxicas e o padrão de seu comportamento. Além disso, destaca-se a identificação e representação das unidades léxicas em pleno contexto de uso, na sua ocorrência nos *corpora*, configurando termos e UFEEs. Tal perspectiva descritiva é também um diferencial do TERMISUL, posto que priorizamos o contexto e preferimos contribuir para harmonizar a terminologia antes que prescrever formas linguísticas preferenciais.

A primeira dificuldade a enfrentar foi a composição do *corpus* textual, fonte da linguagem de uma subárea especializada ainda em fase de consolidação, uma vez que encontramos desigual produção digitalizada correspondente a línguas e países. No português brasileiro, a coleta de comunicações em eventos de profissionais conservadores-restauradores foi numerosa, mas os textos produzidos em instituições universitárias, como monografias, dissertações e teses foram difíceis de encontrar. No entanto, o número de artigos científicos publicados por pesquisadores, manuais e apostilas voltados para o treinamento de funcionários de bibliotecas e arquivos e para a formação de técnicos foi satisfatório. A fim de assegurar que o *corpus* pudesse ser representativo da linguagem especializada em pauta, pesquisou-se intensamente a Internet em busca de diferentes tipologias textuais de modo a atender o público alvo. Os textos coletados tendem a refletir

dois níveis de comunicação: especialista falando com o especialista e o especialista ensinando o aprendiz. Dessa maneira, pensamos contemplar não só os conservadores-restauradores experientes e aqueles ainda em formação nos cursos técnicos e acadêmicos, mas também os tradutores e redatores da área.

A construção do *corpus* PAPEL, datada de 2017, foi projetada segundo a metodologia corrente da Linguística de *Corpus* iniciada pelos pioneiros dos anos 60, confirmada no famoso Simpósio Nobel de 1991 (Svartvik, 1991) e divulgada no Brasil por Berber Sardinha (2004). Um de seus dogmas é a autoria dos textos de falante nativo. No entanto, no mundo globalizado em que vivemos, como atestar a naturalidade de um especialista a não ser por uma análise linguística de seu texto e do seu uso da língua especializada? Por outro lado, como poderá quem não conhece a língua estrangeira conhecer a pesquisa de um conservador-restaurador divulgada em idioma diferente do seu? Esperar pela sua tradução? E o tradutor além de nativo dessa língua será um experiente usuário da língua especializada? Dentre as seis línguas do *corpus* PAPEL, a língua inglesa foi a que encontrou mais material. Deveríamos descartar tais textos porque assinados por pesquisadores e profissionais de nomes latinos, eslavos, semitas e outros? Serão eles anglofalantes nativos? O inglês na contemporaneidade é a língua franca da ciência e, ainda que sua hegemonia cause lamentável detrimento das outras línguas, não podemos negar a realidade.

Diante dessa situação, o critério da autoria de falante nativo foi modulado pelo critério da autenticidade da língua inglesa da fonte da publicação, seja essa um periódico científico com corpo editorial e avaliação das contribuições confiada aos pares da especialidade, seja uma instituição acadêmica reconhecidamente autorizada. Em outras palavras, a autoridade da publicação foi considerada a garantia da aprovação da comissão editorial, banca de examinadores do curso

universitário ou comitê científico do evento e que, portanto, sua linguagem é autêntica.

O aumento considerável de publicações sobre a conservação-restauração disponíveis três anos depois da construção do nosso *corpus*, nos leva a pensar que, hoje em dia, poderíamos ter conseguido muito mais material em todas as línguas. Isso pode ser explicado por várias razões, desde o crescente interesse mundial pela conservação-restauração dos bens culturais, atestado pela a criação de cursos universitários, eventos de associações de pesquisadores e profissionais, até o desenvolvimento tecnológico e a facilidade do seu acesso na *web*. Outro fator, não menos importante, deve ser mencionado, o limitado conhecimento da área da conservação-restauração dos bens culturais em suporte papel de que a equipe dispunha no início do Projeto e que foi gradualmente se aprofundando e ampliando.

A identificação das unidades terminológicas e fraseológicas representativas da linguagem especializada em foco presente no *corpus* ptPP contou com a ajuda de duas ferramentas digitais, AntConc (Anthony, 2014) e SketchEngine (Killgariff *et al.*, 2014). A função *KeyWord* do programa Sketch Engine, tendo como *corpus* de referência o Corpus of the Portuguese Web (Corpus ..., 2011), com os critérios estabelecidos pelo Projeto, gerou uma lista de candidatos a termo. Essa lista apresentou itens que apesar de seus altos índices de chavidade, isto é, embora fossem mais frequentes no *corpus* especializado do que no *corpus* de referência, pareciam não ser específicos da área em foco, mas sim das ciências exatas, ciências biológicas, da saúde entre outras. Neste momento, buscamos a consultoria dos especialistas da área, em particular nos servimos da árvore de domínio de Bojanoski (2018, p. 162-165) a fim de melhor entender como palavras como *luvas e baratas*, conservando seu gênero próximo e diferença

específica da língua comum, podem ser termos da área da conservação-restauração dos bens culturais em formato papel.

Para resolução da questão, ainda recorremos aos conceitos de pertinência temática e de pertinência pragmática (Cabré, 1994; Maciel, 1996). Seleccionamos não apenas termos cuja pertinência temática é indiscutível, ou seja, termos relativos a conceitos próprios da área especializada como *bem cultural*, mas também termos que merecem constar da terminologia *lato sensu* por sua pertinência pragmática, isto é, pela função que desempenham na complementação da informação necessária ao usuário (p. ex.: *valor histórico, valor permanente*). A multidisciplinaridade da área se manifesta na terminologia híbrida de modo marcante e, conforme posteriormente percebemos, a prevalência de determinados conjuntos de termos de outras áreas parece sugerir diferenças de abordagem da problemática da conservação nos *corpora* das outras línguas contempladas pelo Projeto.

Diretamente relacionado com as características dessa linguagem e fator determinante na escolha das fontes de coleta de dados e da decisão do formato do instrumento terminográfico planejado, está o desafio de equacionar a meta de contribuir para a harmonização da terminologia da área e as necessidades do público usuário visado constituído por brasileiros de diferentes níveis de interesse e formação na área da conservação-restauração de bens culturais em papel. No ambiente cibernético de hoje, um instrumento de referência impresso em papel torna-se rapidamente obsoleto e as atualizações periódicas não conseguem acompanhar o fluxo da informação, enquanto o instrumento *on-line* pode ser mantido a par daquilo que está sendo feito na área especializada. No entanto, para que esse instrumento realmente cumpra seu papel, é preciso que sua estrutura acompanhe os avanços da tecnologia contemporânea na estruturação

de bases de dados e não seja apenas um dicionário tradicional amigável e de livre acesso migrado para a mídia digital.

Além disso, é preciso atentar também para o contínuo avanço da terminografia atual que pretende, cada vez mais, dar conta das necessidades dos usuários dos produtos terminográficos. A internet oportuniza o acesso à informação tornando o conhecimento disponível a todos, de tal modo que a linguagem especializada não é mais privilégio dos especialistas da área. Assim sendo, o público alvo de uma base de dados terminológica *on-line* abrange interessados de todos os níveis de especialização, alunos de cursos técnicos e acadêmicos, arquivistas, bibliotecários, museólogos, pesquisadores, cientistas, profissionais, artesãos, e, não menos importantes, mediadores da informação, tradutores, redatores, jornalistas, críticos, marchands e curadores de arte. Não se trata de vulgarização da ciência, da arte e da técnica, mas de democratização do acesso ao conhecimento. Vale ainda lembrar que a avalanche da informação não é monolíngue, pelo contrário, uma de suas características marcantes é o plurilinguismo. Nesse contexto, quem procura a informação não pode esperar que o texto seja traduzido e divulgado. Portanto, para realmente auxiliar o usuário, a base de dados terminológica na contemporaneidade precisa ser ao menos bilíngue, preferencialmente multilíngue. Nesse sentido, a BDT PAPEL busca oferecer, além dos termos e unidades fraseológicas, seus sinônimos e equivalentes em cinco línguas, acessíveis por remissivas indicadas como *hiperlinks*.

CONSIDERAÇÕES FINAIS

No âmbito deste texto, buscamos trazer algumas especificidades da linguagem da área da Conservação e Restauração de Bens em papel principalmente no que tange à sua terminologia e fraseologia.

Nesse sentido, apresentamos os aspectos teóricos e metodológicos que sustentam a elaboração de uma base de dados terminológica multilíngue (BDT PAPEL) dirigida a tradutores, redatores, aprendizes e demais interessados na área e que está disponível *on-line* e de forma gratuita. Destacamos as etapas metodológicas e alguns aspectos que revelam a complexidade de se trabalhar com uma linguagem que está em afirmação e com seis línguas simultaneamente.

Tais aspectos impuseram desafios que fizeram o grupo de pesquisadores buscar soluções que abarcaram aspectos relativos à construção dos *corpora* de estudo, à seleção dos termos e à identificação de suas variantes em português e nas línguas estrangeiras, à identificação das UFEEs e à forma de apresentação dessas informações na base de dados.

Como nos demais produtos terminográficos elaborados pelo Grupo TERMISUL, nosso intuito foi não só identificar as unidades especializadas da área da Conservação-Restauração de Bens em papel, mas sobretudo descrever e representar essas unidades em seus contextos de uso. Essa perspectiva nos fez refletir sobre diferentes aspectos teóricos e aplicados e estabelecer uma correlação entre eles, o que nos permitiu chegar a uma proposição de base que pudesse dar conta, por exemplo, da variação terminológica e da fraseologia.

Os resultados obtidos demonstram que temos feito avanços importantes no que tange à aplicação de uma teoria de caráter linguístico, textual e comunicativo da Terminologia que sustenta a elaboração de produtos que buscam, por um lado, representar as linguagens especializadas e seus usos e, por outro, atender as demandas dos consulentes visados.

REFERÊNCIAS

ANTHONY, Laurence. **AntConc (Version 3.5.8)** [Computer Software]. Tokyo: Waseda University, 2014. Disponível em: <https://www.laurenceanthony.net/software.html>. Acesso em: 21 set. 2021.

ARCOS, Manuela. **Identificação e análise de UFE eventivas na área da conservação e restauração de bens culturais móveis em suporte papel**. 2019. 166 f. Dissertação (Mestrado em Estudos da Linguagem) – Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2019.

BARBOSA, Alessandra Andrade França. **Dicionário Ilustrado de Conservação de Documentos Gráficos**. Dissertação (Mestrado em História Social) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2017. Disponível em: <https://teses.usp.br/teses/disponiveis/8/8138/tde-25092019-144259/pt-br.php>. Acesso em: 13 set. 2021.

BERBER SARDINHA, Tony. **Linguística de Corpus**. Barueri: Manole, 2004.

BEVILACQUA, Cleci Regina. **Unidades Fraseológicas Especializadas Eventivas**: descripción y reglas de formación en el ámbito de la energía solar. 2004. 241 fls. Tese (Doutorado em Linguística Aplicada) – Instituto Universitário de Linguística Aplicada (IULA), Universidade Pompeu de Fabra, Barcelona. 2004. Disponível em: <https://www.tdx.cat/bitstream/handle/10803/7515/tcrb.pdf?sequence=1>. Acesso em: 24 set. 2021.

BIBER, Douglas; CONRAD, Susan; REPPEN, Randi. **Corpus linguistics**: investigating language structure and use. Cambridge: Cambridge University Press, 1998.

BOJANOSKI, Silvana de Fátima. **Terminologia em Conservação de bens culturais em papel**: produção de um glossário para profissionais em formação. 2018. 292 fls. Tese (Doutorado em Memória Social e Patrimônio Cultural) – Programa de Pós-Graduação em Memória Social e Patrimônio, Universidade Federal de Pelotas, Pelotas, 2018.

CABRÉ, María Teresa. A Terminologie et dictionnaires. **Meta**, n. 4, v. 39, p. 590-597, 1994.

CABRÉ, María Teresa. **Terminología**: representación y comunicación. Una teoría de base comunicativa y otros artículos. Barcelona: Universitat Pompeu Fabra, Institut Universitari de Lingüística Aplicada, 1999.

CABRÉ, María Teresa. Sumario de principios que configuran la nueva propuesta teórica y consecuencias metodológicas. In: CABRÉ, María Teresa; FELIU, Judit (Eds.). **La terminología científico-técnica: reconocimiento, análisis y extracción de información formal y semántica**. Barcelona: Universitat Pompeu Fabra, Institut Universitari de Lingüística Aplicada, 2001. p. 17-25.

CÓDIGO DE ÉTICA DO CONSERVADOR-RESTAURADOR. São Paulo, 1996. Disponível em: <https://www.arcoit.com.br/wp/wp-content/uploads/2018/04/CODIGO-DE-ETICA.pdf>. Acesso em: 21 set. 2021.

CORPUS OF THE PORTUGUESE WEB PTTENTEN. 2011. Disponível em: <https://www.sketchengine.eu/pttenten-portuguese-corpus/>. Acesso em: 08 ago. 2021.

FAULSTICH, Enilde. A socioterminologia na comunicação científica e técnica. **Ciência e Cultura**, v. 58, n. 2, p. 27-31, jun. 2006. Disponível em: <http://cienciaecultura.bvs.br/pdf/cic/v58n2/a12v58n2.pdf>. Acesso em: 27 set. 2021.

GAUDIN, François. **Pour une socioterminologie**. Des problèmes semantiques aux pratiques institutionnelles. Rouen: Publications de l'Université de Rouen, 1993.

GAUDIN, François. La socioterminologie. **Langages**, v. 39, n. 157, p. 80-92, mar., 2005.

GUEDES, Maria Tarcila Ferreira; MAIO, Luciana Mourão. Bem Cultural. In: THOMPSON, Analúcia *et al.* **Dicionário IPHAN do Patrimônio Cultural**. Brasília: IPHAN, 2016. Disponível em: <http://portal.iphan.gov.br/dicionarioPatrimonioCultural/detalhes/79/bem-cultural>. Acesso em: 30 ago. 2021.

HURTADO ALBIR, Amparo. **Traducción y traductología**. Madrid: Gredos, 2001.

ICOMOS. **Preservación del Patrimonio Cultural: Hacia un perfil europeo del Conservador-Restaurador**. Pavia: ICOMOS, 1997. Disponível em: <https://icomos.es/wp-content/uploads/2020/01/22.documento-pavia1997.pdf>. Acesso em: 23 set. 2021.

KILLGARIFF, Adam *et al.* **The Sketch Engine: Ten Years On**. 2014. Disponível em: https://www.sketchengine.eu/wp-content/uploads/The_Sketch_Engine_2014.pdf. Acesso em: 23 set. 2021.

KRIEGER, Maria da Graça. Do reconhecimento das terminologias: entre o linguístico e o textual. *In*: ISQUERDO, Aparecida Negri; KRIEGER, Maria da Graça. (Org.). **As ciências do léxico**. Lexicologia, Lexicografia, Terminologia, v. 2. Campo Grande: Ed. UFMS, 2004. p. 327-339.

KRIEGER, Maria da Graça. Porque Lexicografia e Terminologia: relações textuais. *In*: ENCONTRO DO CÍRCULO DE ESTUDOS LINGÜÍSTICOS DO SUL, 8., 2008, Pelotas. **Anais [...]**. Pelotas: Educat, 2008.

KRIEGER, Maria da Graça. **Terminologia Linguístico-Textual (TLT)**: princípios teóricos e metodológicos. Apresentação Oral. Córdoba: Universidad de Córdoba, 2017.

MACIEL, Anna Maria Becker. Pertinência pragmática e nomenclatura de um dicionário terminológico. **Revista Internacional de Língua Portuguesa**, n. 15, p. 69-76, 1996.

MACIEL, Anna Maria B. Terminologia e Corpus. *In*: TAGNIN, Stella; BEVILACQUA, Cleci Regina (Org.). **Corpora na Terminologia**. 1. ed. São Paulo: HUB Editorial, 2013. p. 29-45.

POVEDA, Mireia Xarrié. **El llenguatge de la conservació-restauració d'obres d'art**. 2015. 224 fls. Tese (Doutorado em Didàctica de les Ciències Socials i Patrimoni) – Universidade de Barcelona, Barcelona, 2015. Disponível em: https://diposit.ub.edu/dspace/bitstream/2445/66226/8/MXP_TESI.pdf. Acesso em: 06 ago. 2023.

REDINI, Luana Lara Safar. **Análise do mercado de trabalho do conservador-restaurador de bens móveis e integrados**: uma visão do sistema de preservação a partir da Superintendência do IPHAN no Rio de Janeiro. 2019. 198 fls. Dissertação (Mestrado Profissional) – Instituto do Patrimônio Histórico e Artístico Nacional. Instituto do Patrimônio Histórico e Artístico Nacional, Rio de Janeiro, 2019. Disponível em: http://portal.iphan.gov.br/uploads/ckfinder/arquivos/REDINI_Luana-Dissertacao_Mestrado.pdf. Acesso em: 23 set. 2021.

REGULAMENTAÇÃO DA PROFISSÃO DE CONSERVADOR E RESTAURADOR DE BENS CULTURAIS NO BRASIL. 2013. Disponível em: https://secure.avaaz.org/community_petitions/po/REGULAMENTACAO_DA_PROFISSAO_DE_CONSERVADOR_E_RESTAURADOR_DE_BENS_CULTURAIS_NO_BRASIL/. Acesso em: 17 set. 2020.

TEMMERMAN, Rita. **Towards new ways of terminology description**. Amsterdam: John Benjamins, 2000.

SPINELLI JUNIOR, Jaime. **Guia de Preservação e Segurança da Biblioteca Nacional**. 2009. 112 fls. Dissertação (Mestrado em História Política e Bens Culturais) - Mestrado Profissionalizante em Bens Culturais e Projetos Sociais. Fundação Getúlio Vargas, Centro de Pesquisa e Documentação de História Contemporânea do Brasil, Rio de Janeiro, 2009. Disponível em: <https://bibliotecadigital.fgv.br/dspace/handle/10438/2703>. Acesso em: 21 set. 2021.

SVARTVIK , Jan (Ed.). **Directions in Corpus Linguistics**: proceedings of Nobel Symposium 4-8, 1991. Berlin: Mouton De Gruyter, 1992.

STUBBS, Michael. **Text and corpus analysis**. Computer-assisted studies of language and culture. London: Blackwell, 1996.

STUBBS, Michael. **Words and phrases**: *corpus* studies of lexical semantics. Oxford: Blackwell, 2001.

METALEX: POR UMA METAFERRAMENTA LEXICAL

Guilherme Fromm

INTRODUÇÃO

A escrita deste texto, a partir de um relatório de pós-doutorado, realizado entre 2016 e 2017, deriva do meu trabalho como terminógrafo e professor na área de Ciências do Léxico (em nível de pós-graduação), Língua Inglesa e Tradução (em nível de graduação). Já em minha tese de doutorado (Fromm, 2007), propus a criação de um ambiente computacional voltado para facilitar os trabalhos terminográficos: o VoTec. Após quase quinze anos da pesquisa original, considero o trabalho um tanto datado, tendo em vista, principalmente, a evolução dos programas, ambientes e ferramentas disponibilizados, de forma gratuita ou não, para todos os pesquisadores. Entendo, também, que, embora haja ambientes tão elaborados (computacionalmente falando) quanto o VoTec, já desde aquela época, como o e-Termos, os profissionais que trabalham com o léxico ainda precisam lidar (e, portanto, ter uma considerável curva de aprendizado) com vários programas/ambientes/ferramentas, ao mesmo tempo, para desenvolverem seus projetos.

A ideia desse meta-ambiente lexical (MetaLex), portanto, é agregar, em um só lugar (preferencialmente *online*), todos os tipos de ferramentas e recursos possíveis para o desenvolvimento de trabalhos

em Lexicografia e Terminografia. Não proponho o ambiente em si, especialmente porque não sou da área de Linguística Computacional ou programador e porque acredito que o seu desenvolvimento requereria uma equipe ou uma investigação de doutorado nessa área, mas sim a base teórica para a seleção de todas as ferramentas e/ou recursos que comporiam esse ambiente.

1. BASES TEÓRICAS

É importante destacar que adoto, neste trabalho, a divisão proposta por Barbosa (2001) para os tipos de obras lexicais, como podemos verificar no Quadro 1.

Quadro 1 – Divisão proposta por Barbosa (2001) para obras lexicais

Dicionário	Vocabulário	Glossário
<i>Nível do sistema</i>	<i>Nível da norma</i>	<i>Nível da fala</i>
Trabalha com todo o léxico disponível e o léxico virtual	Trabalha com conjuntos manifestados dentro de um recorte de língua ou de uma área de especialidade	Trabalha com conjuntos manifestados em um determinado texto
Unidade: lexema (significado abrangente; frequência regular)	Unidade: vocábulos/termos (significado restrito; alta frequência)	Unidade: palavras (significado específico; única aparição)
Apresenta (teoricamente) todas as acepções de um mesmo verbete	Apresenta todas as acepções de um verbete dentro de uma área de especialidade	Apresenta uma única acepção do verbete (dentro de um contexto determinado)
Perspectivas: diacrônica, diatópica, diafásica e diastrática	Perspectivas: sincrônica e sinfásica	Perspectivas: sincrônica, sintópica, sinstrática e sinfásica

Fonte: Fromm (2007)

O estudo foi iniciado pela comparação de recursos disponíveis através de pesquisas na Internet, algo como desenvolvido por

Almeida, Aluísio e Teline (2003), Almeida, Aluísio e Oliveira (2006), Fromm (2004), e tantas outras pesquisas e relatórios técnicos desenvolvidos por pesquisadores autônomos e grupos de pesquisa, como o Núcleo Interinstitucional de Linguística Computacional, do ICMC/USP. É através da análise do “estado da arte” que podemos tomar ciência do que existe de mais inovador na área e propor algo além, como um ambiente completo para lexicógrafos e terminógrafos.

Em relação à pesquisa bibliográfica, comecei com obras referentes à lexicologia/lexicografia (mais abrangentes): Atkins, Fillmore e Johnson (2003), Béjoint (2010), Borba (2003), Schmitz (2000), Welker (2004), Xatara, Bevilacqua e Humblé (2011). Na sequência, trabalhei com alguns textos relacionados à terminologia/terminografia e obras de referência na área: Almeida, Pino e Souza (2007), Alves (1996), Aubert (1996), Barbosa (1989, 1991, 1994), Bergeron e Kempa (1995), Bevilacqua e Finatto (2006); Correia (2016); Faria (2014), Finatto (1994, 1998, 2007), Finatto, Lopes e Ciulla (2015), Fromm (2002), Fromm e Yamamoto (2013), ISO 1087, Krieger e Finatto (2004), Perroti-Garcia (2003), Tagnin e Bevilacqua (2013). Também alguns dicionários eletrônicos foram verificados, como Goyos Jr. (2000) e Houaiss (2002).

2. ALGUMAS DEFINIÇÕES

Antes de começar a análise em si, acredito ser interessante propor uma divisão entre os três tipos de análises de programas já existentes que foram realizadas neste estudo: recursos, ferramentas e ambientes. Esses programas podem estar disponíveis em dois meios: para download e instalação no computador, ou como uma página da internet. Destaco que todo o ponto de vista adotado foi como o de

um pesquisador da área, ou seja, todos os programas foram analisados tendo em vista sua utilidade para os lexicógrafos e terminógrafos.

Defino, aqui, **recursos** como programas simples, páginas da Internet ou bases de dados desenvolvidas para um fim e que servem como fonte de pesquisa para estudos lexicais. Bases lexicais (como a WordNet), por exemplo, podem servir para trabalhos nas áreas de Linguística Computacional, de Tradução (automatizada), de Processamento de Língua Natural (PLN) etc. Nem todas foram produzidas exclusivamente para a elaboração de dicionários, vocabulários e glosários, mas podem ser analisadas pelos pesquisadores como subsídios para seus trabalhos. Páginas em que análises de *corpora* podem ser feitas online, através dos grandes portais (como o COCA, o DWDS e AC/DC Linguateca, por exemplo) ou ferramentas de busca que usam a Internet inteira como *corpus* (como o WebCorp), também servem de fonte para checagem de prováveis lexemas, vocábulos/termos e palavras que venham a compor uma obra lexical. Alguns colegas, que têm noções de programação, chegam até a trabalhar com programas (como o SPSS) ou linguagens de programação (como a R) elaborados para trabalhos na área de estatística.

Ferramentas, por outro lado, foram desenvolvidas (as que estamos analisando) para estudos lexicais e/ou linguísticos. As mais antigas são os programas (ou suítes) de análise lexical projetadas para trabalhar com *corpora*, como o WordSmith Tools ou o AntConc, que apresentam três ferramentas básicas: listagem de palavras, listagem de palavras-chave e concordanciador. Embora muitos dessas suítes sejam usadas por várias outras subáreas da Linguística atualmente, os lexicógrafos e terminógrafos continuam a se beneficiar muito dessas ferramentas, já que a base de análise é lexical. Elas disponibilizam os candidatos a lexemas, vocábulos/termos e palavras que comporão a obra de referência, além de apresentarem todo o entorno do item

pesquisado (cotexto, à direita e/ou à esquerda, normalmente disponibilizados em telas KWIC), facilitando a localização de contextos definitórios e/ou explicativos (Aubert, 1996), fornecendo subsídios para a elaboração de uma definição para esses candidatos, por exemplo (como no caso do VoTec). As ferramentas foram sendo aprimoradas com o tempo, muitas migraram para a Internet e algumas até se especializaram unicamente na área do léxico; outras se adaptaram para tarefas específicas que podem (ou não) ser adotadas pelos lexicógrafos/terminógrafos, como concordanciadores simples (ParaConc, TextSTAT), anotadores de *corpora* (UAM *Corpus* Tool), etiquetadores (TreeTagger, TermoSTAT), compiladores de gramáticas em formato computacional (UNITEX) e visualizadores para bases de dados (LexiquePro e Blossary).

Os **ambientes** (no caso desta proposta) são conjuntos de programas voltados especificamente para análise lexical. Neste trabalho, fiz uma diferenciação entre os ambientes para a lexicografia e para a terminografia. Entre os que podem ser instalados no computador, destaco o WeSay e o Language Explorer como ambientes lexicográficos e o SDL MultiTerm e o GesTerm como ambientes terminográficos. Daqueles disponíveis para trabalho online, por um lado, não encontrei nenhum ambiente voltado exclusivamente para a lexicografia; para a terminografia (denominados por Wilkens *et al.*, 2012, como ambientes web de gestão terminológica), por outro lado, descobri vários: VoTec, e-Termos, TermWiki (Plus), Terminus e Corpógrafo. A quantidade de ferramentas e funcionalidades variam bastante de ambiente para ambiente, mas se apresentam como plataformas onde os terminógrafos podem executar grande parte de seus trabalhos.

3. ESTADO DA ARTE

No Quadro 2, apresento uma seleção de vários programas e sites (citados e/ou outros que poderiam ter sido pesquisados), resumindo o que foi apresentado na diferenciação entre recursos, ferramentas e ambientes. Parte da escolha dessa seleção, em específico, vem da minha própria experiência na área; parte está ligada às de indicações de sites (como o de Martin Weisser), que elencam centenas de páginas (especialmente àquelas ligadas às áreas da Linguística Computacional) de ferramentas especializadas.

Percebe-se, nesse quadro, que claramente não há nenhuma ferramenta ou ambiente dedicados apenas à Onomástica. Embora uma pesquisa exaustiva, na Internet, tenha sido realizada, nada foi encontrado à época da pesquisa. Em conversas via e-mail com os colegas que trabalham nessa área dentro do GT de Lexicologia, Lexicografia e Terminologia da ANPOLL, apenas recursos foram apontados.

Quadro 2 – Programas e sites levantados na pesquisa

Computador (instaláveis)	Recursos para Onomástica	Nenhum encontrado
	Ferramentas de Análise Lexical/ <i>Corpus</i>	WordSmith Tools, AntConc, TextSTAT, ParaConc, UNITEX
	Ambientes para Lexicografia	WeSay, SIL Fieldworks Language Explorer
	Ambientes para Terminografia	SDL MultiTerm, GesTerm
Computador – Auxiliares	Recurso de Análise Estatística	SPSS Statistics
	Ferramentas: Anotadores e Etiquetadores de <i>Corpora</i>	UAM <i>Corpus</i> Tool, TreeTagger
	Ferramentas de visualização	LexiquePro

Internet	Recursos - Bases Lexicais	WordNet, WordNet Brasil
	Recursos para Onomástica	Nomes no Brasil (Antroponímia), Namsor (Antroponímia)
	Ferramentas de Compilação e Análise Lexical/ <i>Corpus</i>	Sketch Engine, TermoStat
	Portais para análise de <i>corpora</i> fechados	COCA, Corpus do Português, DWDS, PortalMinas, Linguateca AC/DC, COPATrad
	Ferramentas de visualização	Blossary
	Ambientes para Lexicografia	Nenhum encontrado
	Ambientes para Terminografia	VoTec, E-Termos, TermWiki, Terminus, Corpógrafo

Fonte: elaborado pelo autor

4. ORGANIZANDO AS FONTES

Depois desse primeiro levantamento realizado, escolhi alguns programas elencados no Quadro 2 para o desenvolvimento de mini-projetos a fim de verificar suas funcionalidades. Os programas analisados são apresentados no Quadro 3.

Quadro 3 – Programas analisados

Tipo de programa	Programas	
Ambientes	No computador	Language Explorer, Unitex, WeSay
	Na Internet	VoTec, e-Termos, TermWiki (Pro), Corpógrafo, Terminus
Ferramentas	No computador	WordSmith Tools, AntConc
	Na Internet	Sketch Engine, TermoSTAT

Fonte: elaborado pelo autor

Existem infindáveis possibilidades de classificação de todos esses programas, mas o ponto de vista adotado para todo o estudo tem o lexicógrafo/terminógrafo (tanto o profissional experiente com o manuseio de programas de computador e a metodologia da Linguística de *Corpus*, quanto aqueles com pouca experiência e também os aprendizes) como foco. Tendo isso em vista, foram pensados três macrocampos para analisá-los: as informações gerais sobre os produtos, as informações sobre a possibilidade do uso de *corpus/corpora* em um projeto (quando possível/pertinente) e as informações sobre como estruturar o produto lexicográfico/terminográfico. O Quadro 4 mostra essa primeira subdivisão.

Quadro 4 – Macrocampos da ficha de análise

Ferramenta/Ambiente a ser analisado		
Informações gerais sobre o produto	Informações sobre <i>corpus</i>	Informações das estruturas da obra

Fonte: elaborado pelo autor

5. ELABORANDO UM PROTÓTIPO

Todas essas divisões foram concentradas em uma única planilha base, e as informações buscadas se concentram em caixas de rolagem (com um menu selecionável), de seleção (com um botão) ou de texto (a serem preenchidas com texto). Na análise de cada programa, essa planilha base foi copiada, renomeada e completada com as informações a ele relacionadas. Tenho ciência que este tipo de projeto deveria ser executado como um banco de dados, e não como uma planilha. Devido à minha lacuna de aprendizado para trabalhar com o Access ou programa semelhante, no entanto, não houve como realizá-lo a contento. Para a construção da planilha base, os dados que

deveriam estar disponíveis nas caixas (quando era o caso) foram arrolados em duas planilhas separadas, denominadas Caixas de Seleção – Ferramentas e Caixas de Seleção – Estruturas, como podemos ver no exemplo da Figura 1, e depois ligados aos respectivos campos na planilha base.

Figura 1 – Caixas de Seleção – Estruturas e seus respectivos campos (recorte)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Análise de Estruturas													
2	Macroestrutura	Paradigma Informacional	Paradigma Pragmático	Paradigma de F. Equivalente								Tipologia	Variantes Diassistêmicas	
3	Estrutura Geral	Indeterminado	Indeterminado	Bílingue								Dicionário Enciclopédico	Diacrônicas	
4	Front Matter	Abreviatura	Abonações Literárias	Multilingue								Vocabulário	Diatópicas	
5	Middle Matter	Acronimo	Colocações	Em contraste								Glossário	Diaintegrativas	
6	Back Matter	Alomorfes	Estilo										Dialemáticas	
7	Semasiológico	Categoria Gramatical	Exemplos Corpus									Pedagógico	Diafônicas	
8	Alfabético	Coletivo	Exemplos inventados							Paradigma Mídia		Histórico	Diafrequentivas	
9	Alfabético por agrup	Concreto/Abstrato	Frasesologismos							Árvore Hiperbólica/Grafo		Sincrônico	Diaevolutivas	
10	Onomasiológico	Dominio/Ontologia	Variantes diassistêmicas							Cartograma		Diacrônico	Dianormativas	
11	Thesaurus/Ideologi	Etimologia	Informações Enciclopédicas							Diagrama		Filológico		
12	Analogico	Flexão								Figura		Etimológico		
13	Lógico/Reverso	Forma Equivalente	Paradigma Semântico							Linha do tempo		Atlas Lexical		
14	Descritivo	Forma lematizada	Indeterminado							Nuvem lexical		Toponímico		
15	Normativo	Forma por extenso	Alterador/Alterado							Pronúncia		Antropônimo		
16	Recepção	Gênero	Antonímia							Som		Ortopônico		
17	Produção	Glosa	Causa/Efeito							Vídeo		Ortográfico		
18	Monodirecional	Informações de uso	Classe/Instância							Links para Internet		Sintagmático		
19	Bidirecional	Marca registrada	Co-hiponímia									Construções		
20	Remissivas	Nome Científico	Heteronímia									Colocações		
21		Nome próprio	Hiperonímia							Paradigma Definicional		Frasesologismos		
22		Notas	Hiponímia							Definição		Expressões Idiomáticas		
23		Número	Holonímia							Indeterminada		Provérbios		
24		Qualificador/Classificador	Homonímia							Extensional		Citações		
25		Radical	Meronímia							GPDE/Aristotélica/intensional/Anal		Frases		
26	Composição Interna	Regência Nominal	Parassinonímia							Oracional/Implicação		Paradigmático		
27	Abreviaturas	Regência Verbal	Paronímia							Ostensiva		Sinônimo		
28	Apresentação	Separação de sílabas	Processo/Objetivo							Paráfrase		Antonímico		
29	Bibliografia	Sigla	Produtor/Produzido							Sinonímica		Pictográfico		
30	Chave de Pronúncia	Transcrição Fonética	Sinonímia							Síntese		Gestual		
31	Conjugação de verbos	Variação	Troponímia							Acepções		Rimas		
32	Domínio/Ontologia	Variação de pronúncia								Expressões Idiomáticas		Inverso		
33	Gentílicos	Variações Ortográficas								Formas Derivadas		Família Lexicais		
34	Gráficos									Frasesologismos		Marcado		
35	Gramática									Locuções		Neologismos		
36	Inf. Enciclopédicas									Traços distintivos		Arcaísmos		
37	Mapas											Regionalismos		
38	Nomes próprios											Estranglismos		
39	Préfixo	Usuários										Gírias		
40	Síglas	Público em geral										Jargões		
41	Tabelas	Especialistas										Insultos		
42	Verbetes/Nomenclatur	Aprendizes (especialidade)										Tabus		
43	Verbos irregulares	Alunos do Fundamental										Grupos Sociais		
44		Dicionários tipo 1										Lemas diferenciados		
45		Dicionários tipo 2												
		Caixas de seleção - Ferramentas	Caixa de seleção - Estruturas									Planilha Base	Language Explorer	W:

Fonte: elaborada pelo autor

Todos esses campos foram levantados de três maneiras: minha experiência como lexicógrafo e terminógrafo, a pesquisa bibliográfica aprofundada e os campos disponíveis nos programas analisados. A Figura 2 indica como ficou o projeto gráfico dentro da planilha base no Excel, dividida entre os três macrocampos.

Sequencialmente, no Quadro 5, explico a motivação que me levou ao desenvolvimento de cada caixa; no Quadro 6, apresento uma seleção de todos os campos elencados para esta pesquisa e, na Figura 2, o primeiro exemplo de análise, ainda em formato Excel (os demais encontram-se no anexo, em formato de quadros).

Figura 2 – Ficha de Análise desenvolvida: divisão por macrocampos

Nome da obra/programa:		Desenvolvedor:		Endereço:	
Categoria: Ferramentas		Usuário: Aprendiz		Complexidade: Baixa	
Formatos de Exportação:		Formatos de Importação:		Baseado em Corpus? <input type="checkbox"/>	
CSV DOC HTML LFT		TIBX XLF RTF GDT		Alvo principal: Terminografia	
Vantagens		Desvantagens		Subcategoria:	
Gratuito Interface em português Múltiplas funcionalidades Fácil de instalar Instalável em Windows		Pago Interface em língua estrangeira Poucas funcionalidades Difícil de instalar Não instalável em Windows		Quant. Campos: Indeterminada	
				Ambiente de visualização: Corpus - ferramenta simples	
				Tipo de projeto: Individual Colaborativo	
				Observações	
Ferramentas Corpus		Ferramentas aux. Corpus		Usuários do Produto	
Clusters Colocados Compartilhamento de co Concordanciador Criação de listas a partir c Download de corpus		BiolCat/Extrator de corp. Chargrams Classificação de gêneros Comparação de corpus p Conversor de formatos de Detector de corpus correto		Público em geral Especialistas Aprendizes (especialidade) Alunos do Fundamental Dicionários tipo 1 Dicionários tipo 2	
P. Definicional		P. de F. Equivalente		Paradigma Informacional	
Síntese Anagramas Expressões Idiomáticas Formas Derivadas Frasesologismos Locuções		Bilingue Multilingue Em contraste		Indeterminado Abreviatura Acronimo Alomorfes Categoria Gramatical Coletivo Concreto/Abstrato Dominio/Ontologia Etimologia Flexão Forma Equivalente Forma lematizada Forma por extensão	
P. Semântico		P. Midia		P. Pragmático	
Alterado/Alterado Antonimia Causa/Efeito Classe/Instância Co-Hiponímia Heteronímia Hiperonímia Hiponímia		Árvore Hiperbólica/Grfo Cartograma Diagrama Figura Linha do tempo		Abonações Literárias Colocações Estilo Exemplos Corpus Exemplos inventados Frasesologismos	
				Variantes Dissistemáticas	
				Diantegrativas Diamediais Diatrísticas Diatétricas Diafrequentees Diaevolutivas Dianormativas	
				Macroestrutura Geral	
				Front Matter Middle Matter Back Matter Semasiológico Alfabético Alfabético por agrupamento	
				Macroestrutura Interna	
				Abreviaturas Apresentação Bibliografia Chave de Pronúncia Conjugação de verbos Dominio/Ontologia	

Fonte: elaborada pelo autor

Quadro 5 – Campos da ficha de análise (recorte): motivação e importância

Macro campos	Campos	Subcampos	Motivação/Importância para o desenvolvimento do software
Informações gerais sobre o produto		Categoria	Divisão das três categorias principais propostas para este trabalho.
		Ambiente de trabalho	Verificar se o trabalho pode ser elaborado em uma única máquina, online, ou desenvolvido em ambas as plataformas.
		Alvo principal do programa	Determinar se o programa tem um perfil mais voltado para a Lexicografia, a Terminografia, ambas, ou para a Onomástica.

Macro campos	Campos	Subcampos	Motivação/Importância para o desenvolvimento do software
Informações gerais sobre o produto		Usuário do programa	A ideia é classificar, de acordo com a minha experiência como profissional da área, os possíveis usuários que podem utilizar o programa: os aprendizes, os profissionais ou ambos.

Fonte: elaborado pelo autor

Quadro 6 – Possibilidades de escolhas nos campos da ficha de análise (recorte)

Macro campos	Campos	Subcampos	Indicadores/Arquivos/Ferramentas/Formatos
Informações gerais sobre o produto		Categoria	Ferramenta; Ambiente; Recurso.
		Ambiente de trabalho	Computador; Internet; Computador e Internet.
		Alvo principal do programa	Lexicografia; Terminografia; Lexicografia/ Terminografia; Onomástica.
		Usuário do programa	Aprendiz; Profissional; Misto.
		Usuário do produto	Público em geral; Especialistas; Aprendizes (especialidade); Alunos do Fundamental (Dicionários tipo 1; Dicionários tipo 2; Dicionários tipo 3; Dicionários tipo 4); Tradutores.
		Complexidade de uso	Baixa; Média; Alta.

Fonte: elaborado pelo autor

Tendo em vista a análise até aqui apresentada, elaborei a minha proposta teórica para o MetaLex, conforme o Quadro 7.

Quadro 7 – Propostas para o MetaLex (recorte)

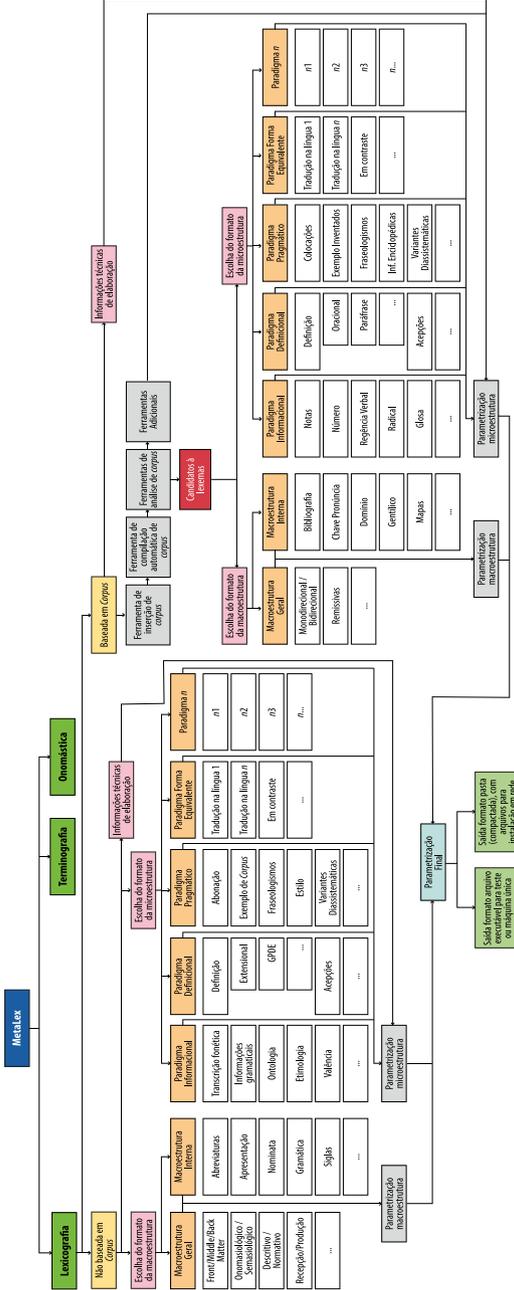
Estrutura	Proposta
Categoria	O programa se configurará como um ambiente na Internet; através de escolhas do usuário (doravante denominado pesquisador máster, PM), campos próprios de análise lexical, ferramentas e recursos poderão ser selecionados, na montagem de um projeto de pesquisa a ser exportado.

Estrutura	Proposta
Ambiente de Trabalho	O trabalho exportado poderá ser instalado na máquina do usuário ou instalado em rede, com acesso liberado a vários usuários.
Alvo principal do programa	O PM, ao começar a usar a plataforma, determinará quais campos comporão o seu projeto. Proponho que, a partir da pesquisa aqui realizada, sugestões de estruturas (macro e micro) e ferramentas (como aquelas ligadas à <i>corpora</i>) sejam indicadas ao usuário, tendo em vista os três grandes tipos de projetos: Lexicográfico, Terminográfico ou Onomástico.

Fonte: elaborado pelo autor

Na sequência (Figura 3), apresento uma proposta gráfica preliminar para desenvolvimento de um projeto no módulo de Lexicografia. A primeira decisão que o denominado PM teria que tomar é se seu trabalho será ou não baseado em *corpus*. Como podemos ver na figura, a estrutura de decisões de um projeto não baseado em *corpus* é um pouco menor do que aquela que se baseia em *corpus*. Na sequência, deverão ser escolhidos os formatos da macroestrutura (formato geral e formato interno, além da escolha de todos os possíveis campos), seguido de sua parametrização, e da microestrutura (seleção dos paradigmas e seus possíveis campos), também seguido de uma parametrização. No caso de projeto baseado em *corpus*, serão carregadas ferramentas básicas (inserção ou compilação automática de *corpus*, análise de *corpus*) e possíveis ferramentas adicionais. A partir da análise do *corpus*, teremos os candidatos a lexemas e, desse ponto em diante, podemos repetir os passos dos projetos sem *corpus*. Nos dois casos, o MetaLex deverá fazer uma parametrização de todos os campos (macro e microestruturas) e ferramentas escolhidas, criar um banco de dados, adicionar as ferramentas selecionadas (para exportá-las), elaborar uma parametrização final e criar a escolha dos processos de saída do projeto: formato executável (que possa ser instalado em um só computador) formato pasta (contendo o banco de dados e as ferramentas escolhidas), compactada, para que possa ser, posteriormente, instalada em um servidor.

Figura 3 – Fluxograma de metaferramenta lexicográfica/terminográfica/onomástica; destaque para a seção lexicográfica



Fonte: elaborado pelo autor

CONSIDERAÇÕES

Embora não considere a execução do trabalho computacional na criação do MetaLex impossível, acredito que o mesmo exigiria centenas (ou até mesmo milhares) de horas de trabalho. Há uma quantidade gigantesca de variáveis já pesquisadas: de ferramentas disponíveis na Internet que teriam que ser adaptadas para a estrutura deste projeto às centenas de campos, cada um com um rótulo e/ou características específicas, que teriam que compor o banco de dados. A experiência que tive com a elaboração do VoTec, quando contratei os serviços da empresa Jr. do ICMC para a elaboração da ferramenta, me ensinou que há um ir e vir constante entre a parte teórica e a prática; muitas das soluções que eu propunha, inclusive, foram incrivelmente facilitadas/resumidas/alteradas a partir de uma percepção diferenciada que os programadores viam em cada desafio.

Por outro lado, um ambiente para trabalhos lexicais desta magnitude, acredito, serviria como uma ferramenta única para os lexicógrafos, terminógrafos, toponimistas e antroponimistas: não haveria mais a necessidade de executar o trabalho em várias fases, já que cada uma delas exige uma ferramenta/ambiente/recurso diferente. O profissional criaria um ambiente de trabalho sob medida para o seu projeto, instalável em seu computador ou num servidor e com capacidade para o trabalho coletivo, evitando ferramentas e recursos desnecessários, contando com tudo aquilo que é essencial para o desenvolvimento de sua obra de referência.

REFERÊNCIAS

ALMEIDA, G. M. B.; ALUÍSIO, S. M.; TELINE, M. F. Extração manual e automática de terminologia: comparando abordagens e critérios. *In:*

WORKSHOP EM TECNOLOGIA DA INFORMAÇÃO E DA LINGUAGEM HUMANA, 2003, São Carlos. **Anais do TIL** [...]. São Carlos: USP, 2003.

ALMEIDA, G. M. de B.; OLIVEIRA, L. H. M. de; ALUISIO, S. M. A terminologia na era da informática. **Cienc. Cult.**, São Paulo, v. 58, n. 2, p. 42-45, jun. 2006. Disponível em: http://cienciaecultura.bvs.br/scielo.php?script=sci_arttext&pid=S0009-67252006000200016&lng=en&nrm=i so. Acesso em: 15 ago. 2021.

ALMEIDA, G. M. B.; PINO, D. H. P.; SOUZA, D. S. L. A definição nos dicionários especializados: proposta metodológica. **RITerm – Debate Terminológico**, n. 3, jan. 2007. Disponível em: <https://seer.ufrgs.br/index.php/riterm/article/view/23812/13814>. Acesso em: 15 abr. 2007.

ALVES, I. M. Definição terminológica: da teoria à prática. **Tradterm**, [S. l.], v. 3, p. 125-136, 1996. Disponível em: <https://www.revistas.usp.br/tradterm/article/view/49898>. Acesso em: 25 ago. 2023.

ATKINS, S. B. T.; FILLMORE, C. J.; JOHNSON, C. R. Lexicographic relevance: selecting information from corpus. **International Journal of Lexicography**, Londres, n. 16.3, p. 251-280, 2003.

AUBERT, F. H. **Introdução à metodologia da pesquisa terminológica bilíngüe**. São Paulo: Humanitas, 1996.

BARBOSA, M. A. Da microestrutura dos vocabulários técnico-científicos. *In*: ENCONTRO NACIONAL DA ANPOLL, 4., 1989. **Anais** [...]. Recife: Anpoll, 1989. p. 567-578.

BARBOSA, M. A. Lexicologia, lexicografia, terminologia, terminografia: objeto, métodos, campos de atuação e de cooperação. *In*: SEMINÁRIO DO GEL, 39., 1991, Franca. **Anais** [...]. Franca: UNIFRAN, 1991. p. 182-189.

BARBOSA, M. A. Da microestrutura de vocabulários técnico-científicos bilíngües para um microsistema terminológico de ecologia e meio ambiente. *In*: SIMPÓSIO RITERM, 4., 1994, Buenos Aires. **Atas** [...]. Buenos Aires: RITerm, 1994. Disponível em: <http://www.riterm.net/actes/4simposio/barbosa.htm>. Acesso em: 11 maio 2007.

BARBOSA, M. A. Dicionário, vocabulário, glossário: concepções. *In*: ALVES, I. M. (Org.). **A constituição da normalização terminológica no Brasil**. São Paulo: FFLCH/CITRAT, 2001.

BÉJOINT, H. **The lexicography of English**. New York; OUP, 2010.

BERGERON, M.; KEMPA, C. **Vocabulaire d'Internet: terminologie des technologies de l'information: vocabulaire anglais-français**. Montreal: Office de la langue française, 1995.

BEVILACQUA, C. R.; FINATTO, M. J. B. Lexicografia e Terminologia: Alguns contrapontos fundamentais. **Alfa**, São Paulo, v. 50, n. 2, p. 43-54, 2006. Disponível em: <http://seer.fclar.unesp.br/alfa/article/view/1410/1111>. Acesso em: 25 mar. 2016.

BORBA, F. da S. **Organização de Dicionários**: Uma introdução à Lexicografia. São Paulo: UNESP, 2003. 356 p.

CORREIA, M. (coord.). **Dicionário de termos do comércio electrónico**. ILTEC (Instituto de Linguística Teórica e Computacional). Disponível em: http://www.iltec.pt/projectos/em_curso/dtce.html. Acesso em: 15 dez. 2016.

FARIAS, V. S. Descrição do léxico em dicionários escolares: proposta para o layout de verbetes de substantivos, adjetivos e verbos. **Travessias**, Cascavel, v. 8, n. 3, p. e10603, 2014. Disponível em: <https://e-revista.unioeste.br/index.php/travessias/article/view/10603>. Acesso em: 15 dez. 2016.

FINATTO, M. J. B. Caracterização de paradigmas definicionais terminológicos. *In*: SIMPÓSIO RITERM, 4., 1994, Buenos Aires. **Atas [...]**. Buenos Aires: RITerm, 1994. Disponível em: <http://www.riterm.net/actes/4simposio/bocorny.htm>. Acesso em: 11 maio 2007.

FINATTO, M. J. B. Microestrutura no dicionário terminológico: bases teóricas e orientações metodológicas em direção ao seu tratamento formal. *In*: SIMPÓSIO RITERM, 6., 1998, Havana. **Atas [...]**. Buenos Aires: RITerm, 1998. Disponível em: <http://www.riterm.net/actes/6simposio/finato.htm>. Acesso em: 11 maio 2007.

FINATTO, M. J. B. Exploração terminológica com apoio informatizado: perspectivas, desafios e limites. *In*: ISQUERDO, A. N.; ALVES, I. M. (org.). **As ciências do léxico**: lexicologia, lexicografia, terminologia. Campo Grande: Editora UFMS; São Paulo: Humanitas, 2007. p. 447-458.

FINATTO, M. J. B.; LOPES, L.; CIULLA, A. Extração automática de candidatos a termos do Curso de Linguística Geral com apoio de recursos da Linguística de *Corpus* e do Processamento de Linguagem Natural. **Domínios de Linguagem**, v. 9, n. 2, 2015. Disponível em: <http://www.seer.ufu.br/index.php/dominiosdelinguagem/article/view/31077>. Acesso em: 05 out. 2021.

FROMM, G. **Proposta para um modelo de glossário de informática para tradutores**. 2002. Dissertação (Mestrado em Linguística) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2002.

FROMM, G. Ferramentas de análise lexical computadorizadas: uma aplicação prática. **Revista Factus**, Taboão da Serra, n. 3, p. 153-164, 2004.

FROMM, G.. **VoTec**: a construção de vocabulários eletrônicos para aprendizes de tradução. 2008. 215 p. Tese (Doutorado em Estudos Lingüísticos e Literários em Inglês) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2008. Disponível em: <https://doi.org/10.11606/t.8.2008.tde-08072008-150855>. Acesso em: 7 out. 2021.

FROMM, G.; YAMAMOTO, M. I. Terminologia, Terminografia, Tradução e Linguística de *corpus*: a criação de um vocabulário bilíngue sobre Linguística. In: TAGNIN, S.; BEVILACQUA, C. R. **Corpora na Terminologia**. São Paulo: Hub Editorial, 2013.

GOYOS JR., D. N. **Dicionário jurídico Noronha**. São Paulo: Observador Legal Editora, 2000. CD-ROM.

HOUAISS, A. **Dicionário eletrônico Houaiss da língua portuguesa**. São Paulo: Objetiva, 2002.

INTERNATIONAL STANDARD ORGANIZATION. **ISO 1087**. Genebra, 1990.

KRIEGER, M. G.; FINATTO, M. J. B. **Introdução à terminologia**: teoria e prática. São Paulo: Contexto, 2004.

PERROTI-GARCIA, A. J. **Vocabulário para ortodontia e ortopedia funcional dos maxilares**: português/inglês/inglês/português. São Paulo: SBS, 2003.

SCHMITZ, J. R. Algumas considerações sobre a dificuldade da utilização de dicionários de aprendizes dentro e fora da sala de aula. In: LEFFA, V. (org.). **As Palavras e a sua Companhia**. Pelotas: Editora Educat, 2000. p. 171-9.

TAGNIN, S. E. O.; BEVILACQUA, C. R. (org.). **Corpora na Terminologia**. São Paulo: Hub Editorial, 2013.

WELKER, A. H. **Dicionários**: uma pequena introdução à lexicografia. Brasília: Thesaurus, 2004.

WILKENS, R. *et al.* Ambientes web de gestão terminológica para criação de produtos terminológicos online. **Debate Terminológico**. n. 8, p. 16-22, jun. 2012. Disponível em: <http://seer.ufrgs.br/index.php/riterm/article/view/29877/18474>. Acesso em: 31 jan. 2014.

XATARA, C.; BEVILACQUA, C. R.; HUMBLÉ, P. R. M. (org.). **Dicionários na teoria e na prática**: como e para quem são feitos. São Paulo: Parábola, 2011.

PROGRAMAS E AMBIENTES ANALISADOS

ANTHONY, L. **AntConc**. Version 3.5.0.w. Disponível em: <http://www.laurenceanthony.net/software/antconc>. Acesso em: 6 out. 2021.

CABRÉ CASTELLVÍ, M. T. **Estació Terminus**. Disponível em: <http://terminus.iula.upf.edu/cgi-bin/terminus2.0/terminus.pl>. Acesso em: 6 out. 2021.

CSOFT. **TermWiki Pro**. Disponível em: <https://pro.termwiki.com>. Acesso em: 6 out. 2021.

DROUIN, P. **TermoSTAT Web**. Disponível em: <http://termostat.ling.umontreal.ca/>. Acesso em: 6 out. 2021.

FROMM, G. **VoTec**. Versão 1.1. Disponível em: <http://treino.votec.ileel.ufu.br/>. Acesso em: 6 out. 2021.

GROSS, M. **UNITEX**. Disponível em: <https://unitexgramlab.org/pt>. Acessado em: 6 out. 2021.

KILLGARRIFF, A. **Sketch Engine**. Disponível em: <https://www.sketchengine.eu>. Acessado em: 30 dez. 2021.

MAIA, B.; SANTOS, D. **Corpógrafo**. Disponível em: https://www.linguateca.pt/corpografo/info_corpografo.html. Acesso em: 6 out. 2021.

OLIVEIRA, L. H. M. **e-Termos**. Disponível em: <https://www.etermos.cnptia.embrapa.br/index.php>. Acesso em: 6 out. 2021.

PAYAP LANGUAGE SOFTWARE DEVELOPMENT. **WeSay**. Disponível em: <https://software.sil.org/wesay>. Acesso em: 6 out. 2021.

SCOTT, M. **WordSmith Tools**. Versão 7. Disponível em: <http://www.lexically.net>. Acesso em: 6 out. 2021.

SIL. **FieldWorks Language Explorer**. Disponível em: <https://software.sil.org/fieldworks/>. Acesso em: 6 out. 2021.

A LINGUAGEM EM TEMPOS DE PANDEMIA: REGISTRO DA TERMINOLOGIA DO NOVO CORONAVÍRUS E DA COVID-19

Márcio Sales Santiago

INTRODUÇÃO

No início de 2020, o mundo foi surpreendido com o surgimento de uma nova doença, que teve origem a partir da mutação do vírus conhecido como SARS-CoV-2. Diante da novidade do vírus e, em consequência da nova doença, que foi denominada pela Organização Mundial da Saúde (OMS) de COVID-19 – termo alfanumérico que significa *Coronavirus Disease 2019* –, muitos termos médicos começaram a circular cotidianamente entre a comunidade científica e a sociedade como um todo, despertando interesse e curiosidade.

No instante em que nos deparamos com a linguagem da Medicina, adentramos em uma nova linguagem, embrenhada de “palavras” e “expressões” desconhecidas. Esse pensamento é bem comum quando se acessam, por meio de textos, tanto de natureza escrita quanto oral, as terminologias específicas de cada área do saber. Isso quer dizer que todo ramo de especialidade tem seus termos, suas definições, suas fra-seologias e sua linguagem especializada.

A Medicina, como uma das mais antigas e sedimentadas áreas do conhecimento humano, desenvolveu uma linguagem hermética e de difícil compreensão, visto que é repleta de particularidades. Rodilla (2005, p. 131) afirma:

Dentro do campo da terminologia científica e técnica, as ciências da saúde e, concretamente, a medicina, apresentam algumas peculiaridades e problemas específicos entre os quais podemos destacar os que derivam dos intercâmbios comunicativos que se estabelecem entre o profissional e o usuário do sistema de saúde ou o leigo em geral.

Vale ressaltar que, em uma época não muito distante, os termos da Medicina, em sua maioria, eram regularmente formados a partir de radicais, prefixos e sufixos gregos e latinos, o que expressa uma tentativa de alcançar o máximo de exatidão no significado, assim como uma troca comunicativa eficiente dos conhecimentos científicos entre cientistas de diferentes línguas e culturas. No entanto, no estudo de Santiago (2007), podemos verificar dois aspectos em relação à terminologia médica:

(1) tal como na língua comum, ou seja, não especializada, a terminologia médica sofre variação tanto no nível do significante, em que se consideram o uso de outras denominações, quanto no nível do significado, em que se consideram ambiguidades e incoerências no âmbito das definições;

(2) observou-se a pouca existência de dicionários e/ou glossários elaborados de forma científica que sejam voltados à consulta da terminologia médica, tanto para a compreensão dos assuntos, quanto para a busca de equivalentes em línguas estrangeiras.

Com relação ao uso da linguagem médica, percebemos que muitas vezes os profissionais utilizam a terminologia padrão propositadamente, como *status* de conhecimento e marca de prestígio. Entretanto, segundo os postulados da Socioterminologia, vertente que reconhece o caráter variacionista que os termos assumem, entendendo que estes têm uma variação sociodiscursiva, não é adequado considerar como “erro” as variações que termos sofrem. Ao agir dessa forma, isto é, negando a variação, o profissional da saúde negará a sua própria língua e todos os elementos socioculturais, já que língua, cultura e sociedade são elementos indissociáveis.

É por conta do hermetismo e da novidade que permeia a linguagem e a terminologia do novo Coronavírus e da COVID-19 que se faz necessário um estudo de seus termos, objetivando a elaboração de um glossário à luz da Terminografia. A finalidade principal desse instrumento será a de auxiliar o consulente no acesso a informações que considere importantes. Dessa forma, em parceria com a literatura médica e um grupo de especialistas, o glossário será mais uma opção que poderá ser utilizada no sentido de compreender o universo terminológico do novo Coronavírus e da COVID-19, pois somente através do entendimento de sua linguagem e de sua terminologia é que teremos condições favoráveis para uma comunicação eficaz.

Por conseguinte, estudar os termos do novo Coronavírus e da COVID-19 pressupõe o compromisso de responder pelo menos duas questões de interesse teórico:

- a. como se constitui e se caracteriza um termo relativo ao domínio enfocado?
- b. de que forma a interferência de outras áreas podem influenciar na constituição da terminologia do novo Coronavírus e da COVID-19?

Um conjunto de fatores interfere na abordagem que se quer realizar em uma pesquisa de natureza terminológica, desde a dimensão, passando pela função, finalidade e metodologia a ser aplicada no estudo. Cumpre ressaltar que a sistematização da terminologia médica e da chamada linguagem popular no plano da saúde contribui para melhorar as condições de informatividade e, conseqüentemente, facilitar a comunicação do profissional de saúde com os pacientes. Acrescenta-se aí um aspecto fundamental relacionado à importância de ferramentas terminográficas como dicionários e glossários, qual seja o de informar sobre os conceitos de uma área, já que todo termo representa um nóculo cognitivo essencial referente ao seu campo do saber.

Neste capítulo, objetivamos apresentar a ferramenta elaborada para o registro dos termos, a ficha terminológica, por ser um dos instrumentos indispensáveis para a elaboração de glossários. Para tanto, faremos a apresentação do panorama teórico da Terminologia, da Terminografia e de seus objetos principais de análise – termo e fraseologia especializada –, os quais são de grande importância para a construção da base metodológica.

1. TERMINOLOGIA E TERMINOGRAFIA: teoria e prática

A preocupação na produção de terminologias é bastante antiga, pois, desde que o ser humano se manifesta e busca conhecimento, encontra-se diante de comunicações especializadas. Como exemplifica Rondeau (1984), os vocábulos especializados já eram utilizados pelas civilizações da Antiguidade Clássica, a citar, a expressão dos pensado-

res e filósofos gregos, a língua de negócios dos comerciantes cretas, a terminologia militar utilizada nas guerras etc.

Por uma questão de padronização, o grego e o latim foram as línguas usadas por estudiosos e pesquisadores nas chamadas nomenclaturas técnico-científicas, cujo papel era o de etiquetar unidades lexicais criadas de forma artificial, no intuito de se evitar fenômenos eminentemente linguísticos como sinonímia e ambiguidades. Temos, portanto, uma classificação científica que a rigor é um sistema organizativo que se rege por um conjunto de regras e critérios que se pretendem universais, mas que dada a grandiosidade do conjunto dos seres vivos e a sua inerente diversidade, sofrem necessárias adaptações. Conforme Santiago (2010, p. 399):

O sistema mais antigo de classificação de seres vivos que se tem conhecimento é o de Aristóteles (384 a.C. – 322 a.C.), que classificou todos os organismos vivos então conhecidos em plantas e animais. Ainda nesse contexto, podemos citar os relevantes trabalhos de Antoine Lavoisier (1743-1794) e Carl Von Linné (1707-1778), ambos no século XVIII, os quais, respectivamente, criaram uma nomenclatura das substâncias químicas semelhante à que ainda está em uso e uma organização da moderna sistemática de classificação para plantas e animais, conhecida atualmente como nomenclatura da Botânica e da Zoologia.

Como outras áreas, a Terminologia possui uma natureza interdisciplinar, pela estreita ligação que mantém com diversas áreas: com a Tradução, na busca do termo preciso durante o trabalho tradutório; com a Linguística, tomando por base as teorias de formação e de estruturação do léxico; com a Lexicologia, baseando-se nos métodos de descrição e de apresentação de informações das palavras; com a Filoso-

fia, na estruturação e na formação de conceitos e conhecimentos; entre outras. Destacamos a afirmação de Sager (1990, p. 4) ao referir que:

A terminologia diz respeito ao estudo e ao uso de sistemas de símbolos e signos linguísticos empregados para a comunicação humana em áreas de atividades de conhecimentos especializados. É primeiramente uma disciplina linguística [...]. Tem caráter interdisciplinar, uma vez que também toma emprestados conceitos e métodos da semiótica, epistemologia, classificação etc. [...] Apesar de a terminologia ter sido no passado muito mais ligada aos aspectos lexicais das línguas de especialidade, o seu escopo abrangia a sintaxe e a fonologia. No seu aspecto aplicado, a terminologia está relacionada à lexicografia e aos usos de técnicas da ciência da informação e da tecnologia.

Com efeito, é relevante, sobretudo nas últimas décadas do século passado, o desenvolvimento da Terminologia, o qual está diretamente associado a duas razões principais: o avanço técnico, científico e tecnológico que experimenta a humanidade e a própria proliferação de termos decorrente de tal avanço. A consequência imediata foi a ocupação e a preocupação por parte dos terminólogos com o surgimento de um elevado número de conceitos e termos novos. Dessa forma, os tipos de comunicações especializadas multiplicam-se em função do nível de formação do público. Com senso crítico inovador, Rey (1979, p. 116) ensina que:

Os vocabulários científicos, técnicos, institucionais, instrumentos obrigatórios da constituição e da transmissão do saber, da harmonização da cultura, do desenvolvimento pedagógico, eram tradicionalmente usados sem serem bem percebidos, salvo pelos próprios especialistas. A tomada de consciência das dimensões linguísticas, formais e, em particular

das léxico-terminológicas dos problemas culturais ou socioeconômicos, torna desejável um grande esforço nesta direção e um desenvolvimento da terminologia.

A ideia de Rey (1979) está na proposição de entender e de tratar a terminologia de característica técnica e científica como pertencente às questões da linguagem e não como algo ideal e homogêneo, que se propõe apenas a uma comunicação restrita a especialistas, desconsiderando fatores linguísticos como sinonímia, polissemia, ambiguidade e variação.

Nesta trilha de abordagem linguística, enfoques como os da Sociolinguística, que trata de uma noção de Terminologia direcionada para o social, considerando a variação como principal ponto que norteia seu desenvolvimento, e o da Teoria Comunicativa da Terminologia desenvolvida por Cabré (1999), colaboraram enormemente para consolidação da área, com objetos de investigação bem delimitados e com princípios teóricos baseados na Linguística, impulsionando novas pesquisas. Logo, o estudo das unidades terminológicas e fraseológicas especializadas passou a ser cada vez mais compreendido como necessário aos diversos campos do conhecimento.

Por sua vez, a Terminografia é definida como o ramo aplicado da Terminologia, o qual se ocupa da elaboração de repertórios terminológicos, tais como dicionários especializados ou glossários, por meio do registro, tratamento e apresentação ordenada de termos de uma dada área técnico-científica. Isso significa dizer que o fazer terminográfico se coloca como uma aplicação das teorizações da Terminologia, que objetiva principalmente elaborar ferramentas como dicionários especializados e glossários temáticos. Segundo Bevilacqua e Finatto (2006, p. 48), a elaboração de um dicionário ou glossário terminológico:

[...] pode ser percebida como um produto imediato, que, tal como o produto lexicográfico, também serve para tirar dúvidas sobre o sentido de um “termo técnico”, em uma área de saber específica. Mas também pode ser visto como produto da reflexão e, ao mesmo tempo, resultado da metodologia derivada dessa reflexão, teoricamente embasada.

Essa linguagem técnica ou científica, segundo nossa ótica, não corresponde a uma língua à parte da língua cotidiana, mas perfaz um uso seu que a torna peculiar, especializada, em uma dada situação de comunicação. O reconhecimento dessa linguagem, procedido segundo determinados princípios e critérios, gerará o dicionário ou glossário.

2. TERMO E FRASEOLOGIA ESPECIALIZADA

O quadro contemporâneo considera que a Terminologia não tem apenas o termo como objeto de estudo, elegendo outros componentes que devem ser considerados na análise terminológica. Observamos, segundo Krieger (2008, p. 1), “uma ampliação no quadro dos objetos de investigação da área tanto que se encontram, por exemplo, proposições que incluem também a fraseologia especializada”. Dessa forma, termos da fraseologia especializada são elementos capazes de transmitir conhecimento especializado, o que os torna componentes essenciais das linguagens especializadas.

No tocante ao termo, trata-se do principal componente linguístico-cognitivo dos textos especializados, constituindo-se, consequentemente, em uma peça-chave de representação e de divulgação do saber científico e tecnológico. Para Benveniste (1989, p. 252):

Uma ciência só começa a existir ou consegue se impor na medida em que faz existir e em que impõe seus conceitos, através de sua denominação. Ela não tem outro meio de estabelecer sua legitimidade senão por especificar seu objeto denominando-o, podendo este constituir uma ordem de fenômenos, um domínio novo ou um modo novo de relação entre certos dados. O aparelhamento mental consiste, em primeiro lugar, de um inventário de termos que arrolam, configuram ou analisam a realidade. Denominar, isto é, criar um conceito, é, ao mesmo tempo, a primeira e última operação de uma ciência.

Desde os primeiros estudos terminológicos, o termo é visto como componente principal da Terminologia. Dessa forma, essa unidade representa o objeto central, mas não único, do estudo terminológico. Segundo Krieger (2001, p. 62), “entender o termo é, de certa forma, entender o sentido maior desta área de conhecimento”.

Com a evolução da Terminologia, a ideia de Wüster (1998) de que o termo é considerado apenas uma unidade cognitiva foi dando espaço a uma nova concepção que conferia a unidade terminológica um caráter linguístico, como bem ilustra Cabré (1993, p. 169):

Os termos, como as palavras do léxico geral, são unidades *sígnicas* distintivas e significativas ao mesmo tempo, que se apresentam de forma natural no discurso especializado. Possuem, pois, uma dimensão sistemática (formal, semântica e funcional) e manifestam também outra dimensão pragmática, uma vez que são unidades usadas na comunicação especializada para designar os ‘objetos’ de uma realidade pré-existente.

Como não existe comunicação especializada sem termos, eles são, portanto, a base de trocas comunicacionais no âmbito das ciên-

cias e das tecnologias, entre tantos outros domínios de interesse da sociedade contemporânea. Daí a importância de se desenvolver estudos que auxiliem a identificá-los e descrevê-los.

Outra definição bastante interessante é a de Gouadec (1990, p. 3), quando afirma que “o termo é uma unidade linguística que designa um conceito, um objeto ou um processo”. Sager (1993), por sua vez, corrobora o pensamento de Gouadec, afirmando que há um conflito entre a necessidade de se denominar e a vontade de normalizar os nomes: “a denominação tem lugar tão pronto como se estabelece um novo conceito, objeto, processo, etc., que inevitavelmente leva a designações desacertadas e a multiplicação de nomes” (Sager, 1993, p. 169).

As afirmações de Gouadec (1990) e Sager (1993) consideram a ideia de que as diferentes áreas do conhecimento humano têm em seus termos a expressão e a construção do saber técnico, científico e tecnológico, além de serem componentes da representação da realidade de um dado campo de saber. Dessa forma, ao ser considerado uma unidade linguística, o termo técnico-científico se comporta como qualquer unidade lexical, podendo, por sua vez, ser passível de mudanças de caráter morfológico, sintático e semântico.

Alpizar Castillo (1997), por sua vez, entende que os termos seguem os tipos de estrutura que o sistema linguístico permite, empregando os mesmos recursos de formação de palavras e submetendo-se às mesmas regras de combinação e suas restrições. Dessa forma, apresenta cinco fatores: função básica que se propõe; temática da qual trata; usuários; situação comunicativa; e situação discursiva.

As concepções descritas revelam verdadeiramente o caráter do termo e sua importância no processo denominativo e nas atividades de conceitualização. Em decorrência de sua natureza de expressar conteúdos específicos, o termo é um componente essencial, nunca

acessório, dos textos especializados, razão pela qual não existe comunicação especializada sem terminologia.

Os termos são, por conseguinte, a base de trocas comunicacionais no âmbito das ciências e das tecnologias, entre tantos outros domínios de interesse da sociedade contemporânea, tal como a Medicina. Daí a importância de desenvolver estudos que auxiliem a identificá-los e registrá-los.

Paralelamente às preocupações com o termo, há cada vez mais interesse sobre as fraseologias especializadas. Por ser um tema complexo e de difícil consenso, muitos pontos de vista teóricos são colocados. No entanto, é fato dizer que os estudos sobre fraseologia se integram às chamadas “Ciências do Léxico” (Biderman, 1998).

Por ser um objeto complexo, multifacetado, de difícil reconhecimento e constituição, possui uma literatura vasta, com variadas denominações, que se diferenciam e se adéquam quanto ao conceito adotado, os quais são divididos em dois grandes grupos: fraseologia da língua comum e fraseologia das linguagens especializadas.

Convém dizer que essa divisão de caráter teórico-metodológico para os estudos referentes à fraseologia faz-se necessária, devido à grandeza e amplitude do tema. Esta separação, como referimos, aloca os estudos acerca das fraseologias pertencentes à língua comum, como *chorar as pitangas*, e fraseologias especializadas, que podemos exemplificar com a construção *fazer febre*, que é um modo de expressão próprio da Medicina e das áreas da saúde afins. A diferença básica entre as duas é que a fraseologia especializada apresenta como um de seus elementos constituintes um termo – *febre* –, considerado como seu núcleo, no caso do exemplo citado anteriormente. Vale ressaltar que consideramos nesta pesquisa as fraseologias especializadas, pois trabalhamos com a terminologia do novo Coronavírus e da COVID-19.

Da mesma forma que o termo, a fraseologia especializada, reflete os modos de dizer de uma área do conhecimento, bem como é considerada uma unidade linguística e comunicativa. Assim se constituindo, torna-se um elemento essencial das linguagens especializadas e das comunicações profissionais.

Em decorrência desse aspecto, tal como o reconhecimento do termo, a identificação das fraseologias especializadas é uma tarefa complexa. Isto porque termos e fraseologias podem ser confundidos, podendo um ser parte integrante do outro. Algumas características dificultam o reconhecimento das fraseologias, como, por exemplo, saber o início e o final da estrutura ou em que medida uma unidade, pelo fato de ser sintagmática, pode ser ou não fraseológica.

3. PROCEDIMENTOS BÁSICOS DA PESQUISA TERMINOLÓGICA

A pesquisa terminológica, para Rondeau (1984, p. 64), está diretamente ligada “ao conjunto das operações de coleta, tratamento e divulgação dos dados terminológicos”. Nesse sentido, tais operações podem ser assim descritas:

- a coleta inclui o recolhimento dos termos, através da análise de textos técnicos ou científicos, do exame de obras lexicográficas especializadas como vocabulários, glossários, artigos ou de obras de divulgação, revistas especializadas, manuais e especificações técnicas;
- o tratamento da terminologia é a totalização de todas as operações para delimitar as noções, identificando o lugar de cada uma delas no mapa conceitual;

- a divulgação está relacionada às operações para colocar à disposição dos consulentes as informações acerca do tema pesquisado na forma de obra de referência, como por exemplo, um glossário.

Para a concretização das operações, é necessário fazermos a constituição do *corpus* de análise, já que a relação produtiva entre *corpus* e Terminografia oferece a possibilidade de coletar, tratar e armazenar um grande volume de informações de uma determinada área do saber por meio de recursos informatizados, que é uma importante ferramenta para o tratamento de dados no contexto das linguagens especializadas. Entretanto, utilizar *corpora* não significa simplesmente usar a tecnologia informatizada para coletar e selecionar termos, desenhar gráficos estatísticos, calcular e comparar listas de frequências. Maciel (2006, p. 6) afirma que “optar pela pesquisa em *corpus* significa, antes de tudo, a adoção de uma visão de língua e uma nova metodologia de trabalho”.

Dessa maneira, a construção do *corpus* exige familiaridade com o domínio a pesquisar, bem como com sua estrutura conceitual, tipologia de textos, organização do discurso e propósitos comunicativos. Ainda de acordo com Maciel (2006, p. 6),

[...] para que os textos coletados se constituam em fonte de informação conceitual, linguística e de uso, além de dominar os procedimentos e técnicas de trabalho, o terminólogo precisa, antes de tudo, saber conciliar as bases teóricas que subjazem à sua proposta e ao tipo de pesquisa em *corpora* que pretende fazer.

Vale dizer ainda que o léxico, seja ele no âmbito geral ou no âmbito especializado, é uma das áreas que mais recebe a atenção dos linguistas de *corpus*, e é também a que mais se beneficia e se projeta

para o mundo. A Linguística de *Corpus*, portanto, postula que a identificação, a análise e a discussão dos dados se desenvolvam a partir de um *corpus* que seja: autêntico, considerando que os usos da linguagem sejam de ordem comunicativa; natural, ou seja, devem conferir os usos de falantes nativos; e criterioso, tendo em vista que o *corpus* seja pertinente à pesquisa desenvolvida (Berber Sardinha, 2004; Biber, 2012).

Logo, seguindo a orientação dos critérios adotados, o *corpus* será constituído de fontes que registram a terminologia referente ao novo Coronavírus e COVID-19, a saber:

- artigos científicos e artigos de divulgação científica sobre a temática enfocada;
- documentos publicados em portais eletrônicos oficiais do Ministério da Saúde do Brasil, das Secretarias Estaduais e de Saúde e da Organização Mundial da Saúde;
- documentos e outros materiais elaborados pela Universidade Federal do Rio Grande do Norte com informações sobre o novo Coronavírus e a COVID-19.

A seleção dos termos e das fraseologias especializadas irá apoiar-se na análise dos textos que constituem o *corpus*. Após essa etapa, passamos para o tratamento da terminologia. O primeiro passo nessa fase é a elaboração das fichas terminológicas.

3.1. A ficha terminológica

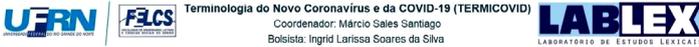
Conforme Cabré (1993), a ficha terminológica deve conter toda a informação que seja relevante sobre um determinado conjunto terminológico. As informações que nelas devem constar são extraídas

dos registros de seleção ou da documentação de referência, de acordo com os critérios estabelecidos previamente pelo terminólogo.

O lançamento dos dados será feito em fichas terminológicas elaboradas em uma base de dados construída no programa *Microsoft Access*. Cumpre salientar que este é um passo indispensável para qualquer trabalho que envolva os procedimentos básicos da pesquisa em Terminologia e em Terminografia Sistemática, pois a ficha serve como uma espécie de controle de qualidade das informações sobre os termos e fraseologias, evitando que não ocorra uma escolha livre ou não sejamos levados a inserir unidades ou expressões que não fazem parte da terminologia enfocada.

Em outras palavras, a grande contribuição da ficha terminológica é a sua função de armazenar as informações que posteriormente serão utilizadas para a elaboração do glossário. A título de ilustração, apresentamos a seguir uma ficha que compõe a base de dados:

Figura 1 – Ficha terminológica da unidade *confinamento*

	
Terminologia do Novo Coronavírus e da COVID-19 (TERMICOVID) Coordenador: Márcio Sales Santiago Bolsista: Ingrid Larissa Soares da Silva	
CÓDIGO 1	
REGISTRO	
Responsável pela inserção	Ingrid Larissa Soares da Silva
Data da inserção	25/10/2021
Revisar	<input checked="" type="checkbox"/>
Responsável pela revisão	Márcio Sales Santiago
Data da revisão	26/10/2021
Concluído	<input type="checkbox"/>
Termo ou fraseologia	confinamento
Variantes em português	[confinamento social]
Categoria gramatical	Nome
Gênero	Masculino
Equivalente em inglês	lockdown, confinement
Equivalente em espanhol	[confinamiento]
Equivalente em catalão	[confinament]
Forma reduzida	
Definição	medida restritiva de prevenção de uma pessoa ou população em espaço limitado por razões de saúde pública, segurança pública ou judicial.
Contexto de uso	O confinamento é importante e, por vezes, essencial para evitar a disseminação do coronavírus e proteger a saúde coletiva, mas são necessárias medidas que reduzam seus efeitos negativos na convivência e no bem-estar. (https://portal.fiocruz.br/iticia/pesquisa-internacional-busca-identificar-o-impacto-social-do-confinamento)
Fonte	FIUCRUZ
Notas	Durante a pandemia da COVID-19, estados e municípios decretaram confinamentos por motivos de saúde pública, sobretudo, para evitar a disseminação do vírus.

Fonte: Projeto TERMICOVID (Santiago, 2020)

Como é possível observar, a ficha terminológica é dividida em três seções. No cabeçalho, estão contidas informações relativas ao projeto, como as instituições e grupos que estão ligados à pesquisa, bem como seus participantes. Logo abaixo do cabeçalho, verifica-se o campo “código”, que serve para ordenar de forma sequencial cada entrada no banco de dados, otimizando a localização futura dos registros.

Em seguida, temos uma seção de registro destinada aos dados de gestão da ficha. Nessa parte da ficha encontram-se as informações sobre quando e por quem o termo foi registrado. No campo “responsável pela inserção”, é possível adicionar o nome do pesquisador responsável pela adição; a cada registro, marca-se o campo “revisar” para que o pesquisador “responsável pela revisão” possa fazer a verificação das informações adicionadas e confirmar ou não se a unidade candidata é um termo ou fraseologia válida e representativa da área temática em questão.

A última seção, reservada ao registro de informações linguísticas e documentais, é constituída pelos seguintes campos:

1. termo ou fraseologia especializada;
2. variantes em português;
3. categoria gramatical;
4. equivalente em inglês
5. equivalente em espanhol;
6. equivalente em catalão;
7. forma reduzida;
8. definição;
9. contexto de uso;
10. fonte;
11. notas.

O campo 1 apresenta o termo ou a fraseologia especializada. Ressaltamos que esses componentes não podem ser encarados de forma isolada, ao custo de ser sempre parte de um conjunto de significados que se relacionam a um determinado domínio especializado.

O campo 2 corresponde às variantes terminológicas ou fraseológicas catalogadas no *corpus*.

No campo 3 é possível o registro de algumas informações gramaticais. Para o modelo de ficha proposto, julgamos pertinente informar a categoria gramatical do termo, podendo ser nome, verbo ou adjetivo; já em relação ao gênero: masculino, feminino, masculino/feminino.

Os campos 4, 5 e 6 são destinados aos equivalentes em língua estrangeira, no caso, em inglês, espanhol e catalão. Quando uma unidade ser caracterizada como empréstimo linguístico, esse campo será descartado no verbete.

O campo 7 é destinado ao registro de informações sobre a forma reduzida, que podem ser do tipo sigla ou acrônimo.

No campo 8 temos a definição, que consiste em apresentar o conceito veiculado pelo termo ou pela fraseologia especializada. Esse paradigma é elaborado de acordo com a estrutura conceitual e o campo temático em que está inserida a unidade.

No campo 9 será registrado o contexto de uso em que ocorre o termo ou a fraseologia especializada.

No campo 10, relativo à fonte, encontra-se a informação referente ao contexto de ocorrência do termo ou da fraseologia registrado no campo 9.

Por fim, o campo 11 indica as notas enciclopédicas ou linguísticas com o objetivo de complementar a informação definitória. São principalmente notas históricas ou fatos relacionados ao termo ou à fraseologia especializada.

CONSIDERAÇÕES FINAIS

À guisa de breves considerações finais, apresentamos pontualmente a proposta de composição do glossário, que deve obedecer aos seguintes princípios e etapas metodológicas:

- o pressuposto de que o termo é um nódulo cognitivo representativo das áreas do conhecimento, ao mesmo tempo que é um componente linguístico e comunicativo;
- o pressuposto de que a fraseologia especializada representa os modos de expressão essenciais de uma área especializada, assim como são estruturas sintagmáticas que contém verbo e termo, apresentado certo grau de fixação;
- a constituição da nomenclatura, no sentido de estabelecer os parâmetros de seleção da terminologia, os graus de especialização da área, a relação da terminologia com o usuário, além da variação terminológica, quando houver;
- a definição dos termos e das fraseologias especializadas indica faces de compreensão de fenômenos e de valores no âmbito das diferentes áreas do saber;

- as definições propostas no glossário terão a função de dar o conceito de cada unidade, atribuindo à obra um caráter didático;
- a terminologia médica comporta variação, o que possibilita disponibilizar outras denominações, que inclui a variação de registro leigo e a sinonímia, além de siglas e acrônimos.

Cumpramos ressaltar que contaremos ainda com o auxílio de especialistas da área da saúde e que tal ajuda será fundamental em nosso trabalho, visto que serão eles quem ajudarão na solução de problemas relativos às definições, por exemplo. Os especialistas desempenharão o papel de consultores do glossário, cabendo a eles a tarefa de validar as definições através de uma atenta leitura dos verbetes.

A organização e divulgação das terminologias, bem como a criação de condições para o registro dos termos, evidencia a importância da organização terminológica. Em razão disso, cada vez mais são criadas obras que compilam e organizam de forma sistemática as terminologias de domínios temáticos especializados. Por meio desses instrumentos de consulta, é possível o acesso, aumento e recuperação da informação destinada a especialistas, tradutores, como também a leigos interessados na temática abordada.

REFERÊNCIAS

- ALPIZAR CASTILLO, Rodolfo. **¿Cómo hacer un diccionario científico técnico?** Buenos Aires: Editorial Memphis, 1997.
- BENVENISTE, Emile. **Problemas de linguística geral II.** Campinas: Pontes, 1989.
- BERBER SARDINHA, Tony. **Linguística de corpus.** Barueri: Manole, 2004.

BEVILACQUA, Cleci Regina; FINATTO, Maria José Borcony. Lexicografia e Terminografia: alguns contrapontos fundamentais. **Revista Alfa**, São Paulo, v. 50, p. 41-52, 2006.

BIBER, Douglas. Representatividade em planejamento de *corpus*. **Cadernos de tradução**, Porto Alegre, n. 30, p. 11-45, 2012.

BIDERMAN, Maria Tereza Camargo. As ciências do léxico. In: OLIVEIRA, Ana Maria Pinto Pires de; ISQUERDO, Aparecida Negri (Orgs.). **As ciências do léxico**: lexicologia, lexicografia e terminologia. v. I. Campo Grande: Editora UFMS, 1998. p. 11-20.

CABRÉ, María Teresa. **La Terminología**: representación y comunicación. Barcelona: IULA/Universitat Pompeu Fabra, 1999.

CABRÉ, María Teresa. **La Terminología**: teoría, metodología, aplicaciones. Barcelona: Antártida/Empúries, 1993.

GOUADEC, Daniel. **Terminologie**: constitution des données. Paris: AFNOR, 1990.

KRIEGER, Maria da Graça. O termo: questionamentos e configurações. In: KRIEGER, Maria da Graça; MACIEL, Anna Maria Becker (Orgs.). **Temas de Terminologia**. Porto Alegre/São Paulo: UFRGS/Humanitas/USP, 2001. p. 62-81.

KRIEGER, Maria da Graça. Terminologia e seus objetos de investigação. In: SIMPOSIO IBEROAMERICANO DE TERMINOLOGÍA: “Terminología, conocimientos, sociedad y poder”, 10., 2008. **Actas [...]**. Montevideo, 2008. p. 1-8. 1 CD-ROM.

MACIEL, Anna Maria Becker. Reflexão sobre a pesquisa terminológica em corpus. In: ENCONTRO NACIONAL DA ANPOLL, 21., 2006, São Paulo. **Anais [...]**. São Paulo: PUC-SP, 2006. p. 1-8. Disponível em <http://www.ufrgs.br/termisul/files/file945269.pdf>. Acesso em: 15 abr. 2018.

REY, Alain. **La terminologie**: noms et notions. Paris: PUF, 1979.

RODILLA, Bertha M. Gutiérrez. La medicina, sus mundos y sus lenguajes. In: CABRÉ, Maria Teresa; BACH, Carme (Orgs.). **Coneixment, llenguatge i discurs especialitzat**. Sèrie monografies. Barcelona: IULA, 2005. p. 131-140.

RONDEAU, Guy. **Introduction à la terminologie**. Québec: Gaëtan Mourin, 1984.

SAGER, Juan C. **A practical course in terminology processing**. Amsterdam: John Benjamins Publishing Company, 1990.

SAGER, Juan C. **Curso práctico sobre el procesamiento de la terminología**. Madrid: Fundación Germán Sánchez Ruipérez/Pirámide, 1993.

SANTIAGO, Márcio Sales. **Terminologia do novo Coronavírus e da COVID-19**: proposta de elaboração de um glossário. Projeto de pesquisa. UFRN, 2020.

SANTIAGO, Márcio Sales. Variação denominativa na terminologia médica: o caso da gripe A H1N1. **Tradterm**, São Paulo, n. 16, p. 397-410, 2010.

SANTIAGO, Márcio Sales. **Redes de palavras-chave para artigos de divulgação científica da Medicina**: uma proposta à luz da Terminologia. 2007. 149 f. Dissertação (Mestrado em Linguística Aplicada) – Programa de Pós-Graduação em Linguística Aplicada, Universidade do Vale do Rio dos Sinos, São Leopoldo, 2007.

WÜSTER, Eugen. **Introducción a la teoría general de la terminología y a la lexicografía terminológica**. Barcelona: IULA/Universitat Pompeu Fabra, 1998.

DESDE QUANDO MEDICAMENTO NÃO É REMÉDIO? ESTUDO COMPUTACIONAL SOBRE O VOCABULÁRIO DE TEXTOS MÉDICOS EM PORTUGUÊS DO SÉCULO XVIII

Maria José Bocorny Finatto

[...] não há uma ‘mágica computacional’ que acontece partindo-se do arquivo de imagem de um livro antigo. Apenas a digitalização acessível [...], por si só, não atenderá todas as necessidades de estudos e pesquisas, ainda que seja um capital hoje indispensável para a conservação do patrimônio histórico documental (Finatto, 2020, p. 26, grifos nossos).

INTRODUÇÃO

Em meio a uma pesquisa de Terminologia Diacrônica (Quaresma; Finatto, 2020), que se estende para a História da Ciência, estamos organizando e explorando um *corpus* de obras médicas antigas em português (Finatto, 2020). Nessa pesquisa, estudamos materiais impressos, na forma de livros, produzidos por médicos, cirurgiões e enfermeiros ao longo do século XVIII. Nosso *corpus* tem sido objeto de experimentos de extração automática de informação textual e terminológica, com técnicas de Processamento

de Linguagem Natural (PLN) e de Linguística de *Corpus* (LC). Trazemos, aqui, em recorte, um conjunto de experimentos que se iniciaram em uma obra de Medicina de 1707, com mais de 600 páginas, escrita pelo médico alentejano João Curvo Semedo (1635-1719). Partindo dessa obra, que comparamos com outras do mesmo período, destacamos alguns resultados sobre:

- a) o emprego e a significação dos termos **medicamento** e **remédio** – e suas variantes de escrita – em contraponto com obras, *corpora* e acervos lexicográficos que incluem o português do século XVIII;
- b) possíveis traços da complexidade ou da acessibilidade desse tipo de texto e discurso.

Neste capítulo, destacamos o papel do exame linguístico-computacional de obras antigas, digitalizadas e transcritas, conforme se faz, hoje, no moderno cenário das *Humanidades Digitais* (Golub; Liu, 2021). As Humanidades Digitais (HD) são um campo de estudos, uma transdisciplina e uma área de atividade acadêmica, situados na interseção da Computação ou das tecnologias digitais e as disciplinas das Humanidades. As HD, assim, incluem o uso sistemático de recursos digitais e informatizados nas Ciências Humanas, bem como análises e ponderações sobre os processos de digitalização.

Nesse destaque das análises linguísticas em meio a algo “super novo” – representado pelas HD –, nossa atenção recairá sobre o vocabulário e as terminologias empregadas ao longo dos textos antigos conforme foram por nós transcritos. A forma peculiar da transcrição dos textos adotada no nosso *corpus* – mais adiante exemplificada – foi estabelecida em meio a um projeto colaborativo de transcrições e análises, iniciado em 2019. Esse projeto visa compor um acervo histórico de natureza didática, dedicado à linguagem médica, reunindo profes-

sores e estudantes de Letras das disciplinas Introdução à Linguística Histórica, Léxico e Dicionários e Terminologia da nossa universidade, a UFRGS (Universidade Federal do Rio Grande do Sul).

Conforme Biderman (2001, p. 19) já nos ensinava, a Terminologia, ao ocupar-se de um subconjunto do léxico de uma língua, percorrendo cada área específica do conhecimento humano, deverá estabelecer relações entre a estrutura conceptual e a estrutura léxica dessa língua. Assim, a partir do exame de textos antigos em português, buscamos encontrar e situar conhecimentos, léxicos e dizeres do passado, mas também queremos enfatizar os desafios, potenciais de aproveitamento e necessidades desse tipo de trabalho. O propósito deste texto, além do relato de experimentos sempre inspirados pela memória da querida autora antes citada, com nossos próprios acertos e erros, é cativar novos pesquisadores interessados nesse tipo de estudo-aventura através do passado.

1. DESAFIOS DO ESTUDO COMPUTACIONAL E LINGUÍSTICO DE TEXTOS ANTIGOS DIGITALIZADOS

Conforme Quaresma, Finatto, Vieira e Trojahn (2022, no prelo), a extração e a categorização de informações oriundas de textos – reunidos em *corpora* ou em coleções digitais – são tarefas importantes do Processamento da Linguagem Natural (PLN). Essa tarefa, em PLN, tornou-se cada vez mais relevante, pois visa estruturar e representar, automaticamente, informações dispersas ao longo de bases de dados, de diferentes dimensões. Assim, o processamento automático da linguagem tem permitido investigar uma grande quantidade de dados em formato digital, de acordo com os mais diversos aspectos próprios de cada área de pesquisa.

Nesse cenário de estudos dedicados à linguagem sob a forma de dados *on-line*, mormente os oriundos da modalidade escrita, encontramos também os acervos de textos antigos, impressos ou manuscritos, de diferentes cronologias e disponíveis em diferentes idiomas. São fontes com enorme potencial de informação e cada vez mais disponíveis em acervos virtuais. Todavia, para os investigadores de PLN, apesar de todos os avanços atuais, como as técnicas conhecidas como *Deep Learning*, Grandes Modelos de Língua (LLM) e Mineração de Textos, explorar esses materiais antigos ainda significa enfrentar uma série de dificuldades. Entre vários desafios, há que se lidar, por exemplo, com a variação ortográfica e tipográfica de materiais pior ou melhor digitalizados, transcritos ou não, com deficiências dos sistemas de OCR (reconhecimento óptico de caracteres), com as diferentes padrões e metodologias de edição de textos antigos, com a falta de glossários e de léxicos computacionais para identificação de formas variantes, entre outras carências. Entre essas, destaca-se a carência de subsídios para se dar conta da transformação histórica dos sentidos de palavras em diferentes contextos e épocas, o que é vital para apoiar a análise semântica automática – ou subjetivamente humana – desse tipo de fontes. Ainda que se tenham ótimos dicionários históricos, há muito a desenvolver, principalmente no âmbito do léxico terminológico da língua portuguesa.

São muitos os desafios do trabalho linguístico-computacional, mais ou menos automatizado, com *corpora* textuais mais ou menos antigos digitalizados. Em meio a eles, o papel do analista linguista e do filólogo adquire uma importância acentuada, sobretudo quando integram equipes de pesquisa coordenadas por cientistas da Computação. Essas equipes, em geral já bastante acostumadas a tratar, com sucesso, materiais escritos da atualidade, encontram uma série de “vazios” que será preciso preencher para o trabalho automático avançar. Em mui-

tos casos, adaptam-se funcionalidades de ferramentas que foram criadas apenas para dar conta de línguas, textos e dados da modernidade.

Nessa linha, a representação automática de características lexicais e semânticas de textos, obras ou de *corpora* antigos dependerá da localização, de um dado discurso, em meio a todo um *frame* semântico sócio-histórico. Esse processo permitirá depreender como e por que se observa uma dada enunciação em um dado tempo, contexto cultural, ideológico e epistemológico. Essa localização, que envolve todo um esforço multidisciplinar e exige insumos específicos, implicará traçar um cenário para o material linguístico sob exame, produzido em um dado espaço-tempo, em uma dada língua-cultura, ambiência que já havíamos denominado como *entorno de significação* (Finatto, 2001).

Portanto, podemos dizer que, mesmo quando disponíveis em formato digitalizado e já mesmo em forma de arquivos de texto, para além de um documento “computacionalmente” visto como uma sequência de caracteres, de palavras e de frases, de *word embeddings* ou de vetores de palavras, o tratamento do conteúdo semântico de materiais antigos – sobretudo em português – segue sendo uma tarefa nada trivial. Além disso, há, via de regra, muito mais recursos para o processamento de línguas históricas como o inglês, e bem menos para outras, como o português. Isto é, há menos recursos para línguas “menos diretamente processadas”, línguas que não têm recursos perpassados por algo que provenha ou se dirija para o inglês. Em estudos diacrônicos com o português, tais carências tendem a desenhar uma espiral de dificuldades. E, trabalhando-se com o português e/ou outras línguas ou variantes “minoritárias”, enfrenta-se uma condição, denominada em inglês, *low-resource language*. Por isso, quando não há o que usar, torna-se praxe verificar como e o quanto alguns recursos e métodos pré-existentes poderiam ser aproveitados ou “reciclados” ou

“traduzidos” para lidarem com *corpora* antigos. Nessa via, será preciso ficar atento aos traços de um “recondicionamento” de funcionalidades arquitetadas e ajustadas para dar conta de características gramaticais de uma língua em particular, que nem sempre se encaixam na nossa língua, ainda mais se a mirada for diacrônica.

Por fim, o desprestígio dos estudos históricos em geral, o menosprezo para os olhares voltados para o passado, algumas vezes vistos como nada prioritários frente à emergências do presente, tendem a relegar ao pesquisador apenas migalhas de financiamento e/ou fomento à pesquisa. E, muitas vezes, especialmente no caso de pesquisas linguísticas, a falta de financiamento nos obriga ao improvisado das ferramentas e das técnicas de trabalho: onde deveria haver um pente, temos um garfo adaptado para a função de pentear. Nesse contexto, apresentar iniciativas de pesquisa linguístico-histórica com apoio tecnológico e/ou computacional avançado pode parecer quimera. De certo modo, é mesmo. Mas, a despeito disso, cremos que valerá a pena seguir a leitura deste texto, pois se pretende demonstrar, justamente, os ganhos do enfrentamento de dificuldades e compartilhar alguns improvisos que, com boa vontade, poderiam até ser considerados inovações.

2. MATERIAL DE ESTUDO EM CONTEXTO, SUPOSIÇÕES DE PARTIDA E MÉTODOS

Como linguistas e terminólogos, no cenário dos estudos de Terminologia Diacrônica ou Sócio-histórica (Dury; Picton, 2009) ou da Teoria Sócio-Cognitiva da Terminologia (Temmermann, 2000), temos lidado com textos médicos antigos. Esses são estudos que, via de regra, na seara dos textos especializados, com temas técnico-cien-

tíficos, enlaçam-se também à História dos Conceitos (Jasmin, 2005; Kabatek, 2005) e à História da Ciência. A Terminologia Diacrônica é uma confluência de pesquisa linguística multidisciplinar hoje revigorada pelas HD.

Nesse cenário, já estavam posicionadas, há tempo, a Filologia Digital (Paixão de Sousa; Trippel, 2006), a Linguística Histórica e a História, compartilhando objetos e focos de estudo ampliados pelos recursos computacionais da Linguística de *Corpus* (LC) (Berber Sardinha, 2004). No contexto ampliado dos estudos linguístico-históricos, já contávamos com trabalhos em diacronia também no campo da Terminologia (Marengo, 2016) e da Lexicografia (Murakawa, 2014). Assim, desenha-se uma trajetória, no qual o trabalho de Santos, Olival e Sequeira (2020), nos dias de hoje, aponta uma série de novas necessidades multidisciplinares. Afinal, em torno de acervos e coleções antigas digitalizados, interagem historiadores, filólogos, linguistas, cientistas da computação, arquivistas, filósofos e outros tantos estudiosos. Todos mirando um mesmo registro linguístico de partida que se transformará em diferentes objetos de ciência.

Nesse contexto digital e multifacetado (Finatto, 2020), na parte que nos cabe, o propósito é descrever e sistematizar textos, discursos e conteúdos de uma série de obras especializadas sobre doenças e seus tratamentos do século XVIII. E, em uma necessária ampliação de escopo instigada pelo “espírito comparativo entre *corpora*” da LC, acabamos estendendo o olhar para obras em espanhol e em francês, destacando as que fossem traduções, fontes ou adaptações coevas dos nossos materiais em português. Para tanto, o suporte computacional tem sido indispensável, assim como o apoio de diferentes especialistas que se dedicam à História em geral, à História da Ciência e da Medicina.

Uma das nossas fontes em estudo, aqui em destaque, é o manual *Observações medicas doutrinaes de cem casos gravíssimos, que em serviço da pátria, & das nações estranhas escreve em língua portugueza, & latina* (Semedo, 1707). Um fac-símile dessa obra está disponível na plataforma *Google Books*, mas também é acessível em diferentes bibliotecas digitais – que usaram volumes físicos de proveniências diferentes, conforme se pode conferir pelos diferentes recursos (e até manchas) de imagem oferecidos. Com essa obra, João Curvo Semedo (1635-1719) dava continuidade a uma peculiar “escola de pensamento” de Medicina associada à sua figura, iniciada em 1680. Sem sermos historiadores, arriscamos dizer que nos parece uma escola alicerçada sobre alguma empiria, caracterizada pela retomada de relatos de casos atendidos, vivências, experiências, ganhos e derrotas em diversos tratamentos e procedimentos. Nesses relatos, marcados por uma série de condições, ignorâncias, dogmas e desconhecimentos científicos, Semedo destaca as figuras dos diferentes pacientes – desde crianças até pessoas idosas. Pessoas que o autor se preocupava em identificar e caracterizar em pormenores bastante pessoais e íntimos, algumas até com o endereço de moradia.

Os atendimentos narrados por Semedo incluíam nobres, religiosos, escravizados, pessoas do povo, trabalhadores, “pessoas de má índole”, mulheres, homens e crianças. São, assim, 101 histórias de sucessos e também de perdas, com muitos episódios que envolvem as suas então famosas invenções farmacológicas. Invenções que ele, sem pudor, recomenda a compra. Para mais detalhes acerca de João Curvo Semedo (1635-1719), recomendamos o trabalho histórico de Lourenço (2016). A Figura 1, a seguir, serve como uma mínima amostra da obra em estudo e também destaca o papel simbólico da imagem do seu autor.

Figura 1 – Observações Médicas e Doutrinais



Fonte: Semedo (1707).

O século XVIII tende a ser reconhecido como um período em que se projetava a realização de um novo mundo inspirado pelo Renascimento, desenhando-se a consolidação e a ampliação, via maior divulgação, nas línguas nacionais, de muitas das ciências por nós hoje conhecidas (Barbosa, 2020). Quanto à linguagem, podemos supor que uma das características desse projeto Iluminista, além do abandono dos registros científicos feitos somente em Latim, residisse também no deslocamento de padrões ou de convenções de uma escrita, ainda barroca, rumando-se a um estilo de expressão mais compatível com os ideais de acessibilidade e de inclusão.

Junto aos ideais de disseminação dos conhecimentos, com a razão sendo contraposta ao pensamento religioso, há todo um conjunto de efeitos potenciais sobre os modos de apresentar as escritas científicas. Como no caso desta obra de Semedo (1707), vemos desejos ex-

pressos, textualmente declarados, de que os escritos sobre os conhecimentos e as artes médicas pudessem ser compreensíveis também por pessoas de menor erudição (Finatto; Quaresma; Gonçalves, 2018).

A citada condição de um leitor de “menor erudição”, muitas vezes, parece resumir-se à presunção de uma pessoa com dificuldades com o Latim. Dificuldades que Semedo buscava elucidar com remissões e notas apresentadas nas laterais de página. Todavia, no corpo do texto da página impressa, menções a obras em Latim correspondem a comentários numerados que remetem aos trechos originais, na eventualidade de o leitor saber Latim e querer conferir a informação traduzida.

Assim, se Semedo, em sua escrita, dialogava principalmente com jovens médicos, cirurgiões e aprendizes, ao mesmo tempo também convoca, para a “conversa”, toda uma série de outras obras latinas, em uma condição que reconhecemos como intertextualidade. Nesse jogo entre a sua fala, em tese facilitada, e as remissões, no desenrolar dos relatos, Semedo, via de regra, destaca – e auto elogia – o valor de seus conhecimentos e experiências. Tudo isso serve para caracterizar seu *modus dicendi*, conformando um estilo que o autor considerava ser acessível também para barbeiros, parteiras, enfermeiros e quaisquer pessoas que, sabendo ler, atendessem doentes em diferentes situações.

Esse comportamento de Semedo não é algo isolado, pois, conforme mencionado, integra um projeto coletivo de popularização e/ou vulgarização de conhecimentos. Em diferentes obras médicas, impressas ao longo do século XVIII, em diferentes países e/ou reinos, encontramos críticas contra o rebuscamento e o hermetismo dos escritos da Ciência. Críticas assentadas na percepção da carência de condições de entendimento e/ou de formação dos diferentes ti-

pos de profissionais e dos “práticos” da Saúde atuantes naquela época (Campos, 2016), como parteiras e barbeiros. Essas críticas também reconheciam a necessidade de haver maior acessibilidade, inteligibilidade e formação adequada para mais pessoas.

Outro exemplo desse esforço, ainda que mais tardio, temos na obra em português intitulada *Aviso à gente do mar sobre a sua saúde* (Mauran; Carvalho, 1794). Nesse livro, encontra-se, textualmente, que o autor – e o seu tradutor português – ambos experientes cirurgiões de armadas nacionais – teriam realizado um “sacrifício de estilo” em prol da facilitação da sua mensagem. Buscando auxiliar os cirurgiões embarcados (ou os marujos que atendiam doentes), a obra menciona a intenção de livrar-lhes dos hermetismos de uma erudição vazia. Nesse texto, vemos que enfrentavam, com a escrita em português, também a pouca familiaridade desses leitores com outras línguas e culturas de escrita. Para mais elementos sobre a formação diferenciada de cirurgiões, médicos e outros profissionais da Saúde no século XVIII, o que não poderemos desenvolver aqui, recomendamos o excelente trabalho de Palma (2021).

Voltando à figura do médico Semedo e às suas obras, poder-se-ia ponderar que propagandear a necessidade de uma linguagem mais simples ou acessível talvez tenha sido apenas um exercício “populista” de sua retórica. Afinal, como dizemos hoje, uma coisa é dizer, outra coisa é fazer. À primeira vista, desde uma mirada que parte da nossa atualidade do século XXI, seus relatos ainda poderiam parecer ser bastante complexos se comparados, por exemplo, a obras como a antes citada, dirigida aos cirurgiões embarcados. Entretanto, conforme vemos, isso é mais um elemento a ser verificado, empiricamente, via exame do texto, em uma exploração que pode ser apoiada por técnicas computacionais.

Quanto ao emprego pontual dos termos *medicamento* e *remédio*, sendo o primeiro de origem mais erudita e, por extensão, em tese, “mais científico”, quisemos verificar o emprego de ambos ao longo das *Observações*. Essa dupla de termos, em seus diferentes empregos e alternâncias, conforme supomos, daria alguma medida entre maior ou menor erudição expressa. Também poderia fornecer pistas sobre uma eventual facilitação ou popularização da informação, ainda que fosse necessário uma ponderação mais ampla sobre o contexto sócio-histórico e lexicológico associada às enfermidades e aos seus tratamentos naquele período (como vemos, por exemplo, em Murakawa, 2014 ou em Palma, 2021). Com essa suposição de partida, munidos de ferramentas, fomos aos textos, conforme o relato que segue.

3. FERRAMENTAS E AMOSTRAS TEXTUAIS SOB EXAME

Para nos ajudar a identificar características associadas ao uso dos dois termos antes citados – *medicamento* e *remédio* – e localizar elementos indicativos de um *continuum* de complexidade desses documentos antigos, utilizamos duas ferramentas computacionais e um recurso lexicográfico combinado a um *corpus* documental.

A primeira ferramenta é um recurso já bastante reconhecido e utilizado em estudos de Linguística de *Corpus* dedicados ao léxico, o AntConc (Anthony, 2018). O AntConc é um sistema que, entre outras funcionalidades, percorre todo um arquivo de texto – ou vários arquivos ao mesmo tempo – e fornece listas de palavras em contextos acompanhadas por dados de frequência e distribuição. Assim, temos dados também sobre combinatórias dos termos em estudo, *medicamento* e *remédio*. Cabe registrar que a noção de *palavra*, nesta ferra-

menta, é a de palavra gráfica, entendida como um conjunto de caracteres entre dois espaços em branco.

A segunda ferramenta a que recorreremos é o sistema NILC-Metrix 3.0 (Gazzola; Leal; Aluísio, 2019), fruto de mais de uma década de trabalho em torno do processamento do português do Brasil. Esse sistema processa um arquivo de texto e o transforma em uma série de métricas e contagens diversas. São gerados dados numéricos que podem subsidiar uma detalhada avaliação linguística da coesão, da coerência e do nível de complexidade do texto em exame. Essas contagens – em número de 200 – incidem sobre características do léxico, da morfologia lexical, da semântica e da sintaxe do texto como um todo. Cabe salientar que esse sistema foi configurado conforme características do português brasileiro da atualidade, apreendidas em *corpora*, bases de dados, dicionários e gramáticas.

A terceira “ferramenta” utilizada não é um recurso computacional propriamente dito. Trata-se de uma obra de referência de suma importância, o *Dicionário Histórico do Português do Brasil: séculos XVI, XVII e XVIII – DHPB* (Biderman; Murakawa, 2021). Junto a essa obra hercúlea, está disponível também o *corpus* documental (CORPUS – DHPB, 2021) especialmente construído para dar suporte ao reconhecimento de palavras e de respectivas significações ao longo do período em questão. Nesses valiosos *corpus e dicionário*, encontramos também importantes fontes documentais sobre temas científicos e médicos.

Para um ensaio com o texto de Semedo (1707), fixamos um *corpus* de estudo amostral dessa obra. Examinamos, preliminarmente, um conjunto composto pela transcrição de 89 dos 101 relatos da obra. Nessa amostra, identificamos, com o AntConc (Anthony, 2014), um universo total de 146.905 itens/palavras (*word tokens*) e 17.144 formas distintas de palavras (*word types*).

3.1. Textos em contraponto

Para contraponto com materiais igualmente antigos, lançamos mão das *Gazetas Manuscritas da Biblioteca Pública de Évora* (Lisboa; Miranda; Olival, 2002) que poderia ser considerado um conjunto de textos de estilo “pré-jornalístico”, produzidos por pessoas leigas, em Portugal, entre 1729 a 1731. Esses materiais, tratando de temas do cotidiano e dirigidos às pessoas comuns do povo da época, pessoas que sabiam ler, são um contraste de “simplicidade” frente ao texto do médico, considerado especializado.

O primeiro volume transcrito dessas *Gazetas* – com o qual trabalhamos – encontra-se acessível no *corpus* histórico Tycho Brahe (<https://www.tycho.iel.unicamp.br/home>). Veja o item *Parsed Corpus of Historical Portuguese*. Um alerta importante a se fazer é que esse *corpus* das *Gazetas* (1729-1731) é apresentado no acervo Tycho Brahe em arquivos identificados como sendo sua *versão modernizada* em *versão [original] transcrita*. Processamos as duas versões, moderna e antiga, mas destacamos aqui, em um quadro, apenas os resultados da versão modernizada no quesito *observação de traços complexidade*. Veja-se a diferença nos segmentos abaixo, a começar por *somana/semana*:

O Marques de Abrantes ainda não tomou somana, e fes em sua caza hua conferencia academica com muitos refrescos. Apareceu na parede de hua igreja de Trancozo hum sepulcro com hum cadaver organizado, e hum livro emque o parrocho Pascoal João no anno de 1533. jurava que era o original de Gonçallo Annes Bandarra comnovos sonhos verdadeiros, e hua copla que verificava este descubrimento; porem ao livro recolheu hum comissario do Santo Officio o qual tribunal tem feito muitas prizões fora de Lisboa.

[15] O Marques de Abrantes ainda não tomou semana , e fez em sua casa uma conferência acadêmica com muitos refrescos. [16] Apareceu na parede de uma igreja de Trancoso um sepulcro com um cadáver organizado, e um livro em que o pároco Pascoal João no ano de 1533 jurava que era o original de Gonçalo Anes Bandarra com novos sonhos verdadeiros, e uma copla que verificava este descobrimento; porém ao livro recolheu um comissário do Santo Ofício o qual tribunal tem feito muitas prisões fora de Lisboa. Fonte: *Corpus histórico Tycho Brahe*. Disponível em: <https://www.tycho.iel.unicamp.br/home> Acesso em: 22 ago. 2023.

Para um contraste inicial com textos da atualidade do século XXI, quanto a um grau de complexidade médio, utilizamos um conjunto de Sentenças de juizados especiais do Tribunal de Justiça do Estado do Rio Grande do Sul – conhecidos como “tribunais de pequenas causas”. Esses textos são reconhecidos como bastante complexos em termos de tessitura sintático-semântica e também quanto ao emprego de terminologias (Motta, 2018; 2021). Abaixo um pequeno trecho de uma dessas Sentenças, produzida em 2016:

Defende a Ré a necessidade de suspensão do trâmite da presente demanda, em razão da existência de processo de recuperação judicial, com antecipação de tutela deferida. Inicialmente, não se desconhece que está em processamento pedido de recuperação judicial da sociedade empresária impugnante, bem como que há recomendação, ressalvado entendimento jurisdicional diverso, no sentido de suspensão dos feitos executórios, em que a referida empresa figure como demandada. Todavia, saliento, em primeiro lugar, que não houve informação, a este Juízo, sobre o efetivo deferimento da recuperação judicial, tampouco do plano respectivo,

com a inclusão dos credores destes autos naquele rol, de modo a se saber quais os prazos de cumprimento das obrigações assumidas, mormente aquelas decorrentes de demandas judiciais oriundas de danos ao consumidor (Motta, 2021, Anexos, *corpus* de estudo. Disponível em: <https://github.com/search?q=Ester+Motta&type=repositories>. Acesso em: 22 ago. 2023).

3.2. Remédios e medicamentos – hoje

Na atualidade, *remédio* e *medicamento* podem parecer, muitas vezes, uma mesma coisa, em uma relação de sinonímia. Essa condição se reforça nas propagandas brasileiras da atualidade, quando as duas formas se alternam ao longo de uma mesma peça publicitária. Também em notícias de jornais, ocorre a alternância *medicamento/remédio*, que parece, *prima facie*, aleatória. Entretanto, em um âmbito técnico especializado, há diferenças importantes entre os traços conceituais associados aos dois termos. Conforme vemos resumido e explicado, por exemplo, no *site* da ABEM (Associação Brasileira de Esclerose Múltipla), entidade que tem como um de seus objetivos esclarecer pessoas leigas sobre pontos importantes associados à doença, temos o seguinte conjunto de informações sobre o distinto uso e valor dos dois termos. A instrução traz a validação em fontes institucionais de Saúde Pública e há, inclusive, autoria assinada por uma profissional especializada:

- **Remédio:** um termo mais amplo, vem do latim *remedium* (cura): Métodos ou cuidados terapêuticos que ajudam a aliviar desconfortos, mas que não passaram pelas etapas que um medicamento passa para ser liberado.

- **Medicamento:** é um produto que foi desenvolvido por uma indústria farmacêutica ou farmácia magistral e que cumpriu todas as exigências legais definidas por órgãos regulatórios. Os efeitos são conhecidos e foram comprovados cientificamente por inúmeros testes.
- **Fitoterápicos e homeopáticos** também são **medicamentos**. Os fitoterápicos são obtidos exclusivamente de plantas medicinais. Já os homeopáticos são derivados de plantas, animais, minerais, de substâncias biológicas ou sintéticas e são utilizados conforme os princípios da Homeopatia.

Medicamentos são produtos que servem para:

- Prevenir o aparecimento de doenças (por exemplo, vacinas).
- Aliviar sintomas ou sinais (por exemplo, medicamentos contra dor e febre).
- Controlar doenças crônicas e reduzir o risco de complicações (por exemplo, medicamentos para pressão alta, diabetes, asma, entre outros).
- Recuperar a saúde (por exemplo, antibióticos).
- Auxiliar no diagnóstico de doenças (por exemplo, contrastes utilizados em radiologia e outros exames).

Enfim, todo medicamento pode ser considerado um remédio, mas nem todo remédio pode ser considerado um medicamento.

Fonte: <https://www.abem.org.br/a-diferenca-entre-remedio-e-medicamento/>, Acesso em: 15 out. 2021, grifos nossos. Autoria creditada à Enfermeira Dra. Ivone Regina Fernandes - Coren-SP 34.495.

4. SUPOSIÇÕES E BUSCAS

4.1. Terminologias em alternância

Já se percebia a emergência e a consolidação da Farmacologia em português, ao longo do século XVIII, por exemplo, na obra *Pharmacopeia Geral para o Reino, e Dominios de Portugal*, de Francisco Tavares, editada em dois volumes¹ em 1794. Muito antes dessa consolidação, desenhava-se a transformação das Boticas em Farmácias, cujo diferencial era contarem com alguma base científica. Nesse percurso, em 1707, temos o registro dos esforços de Semedo por firmar-se como “genial inventor” e médico-farmacologista. Esse médico apresentava-se como um comerciante autorizado por seu notório saber, via de regra divulgando – em suas obras - uma série de tratamentos, de medicamentos (ou remédios?) por ele inventados e testados.

Nossa suposição, em função disso, é que os termos *medicamento* e *remédio* fossem utilizados de modo diferenciado e até paradoxal, na sua obra em foco (1707). Assim, imaginamos que haveria, nas *Observações* de Semedo:

a) algum esforço por aproximar ambos os termos, mesmo que um fármaco “científico”, de sua lavra, não fosse exatamente um *remédio*, pois seus interesses comerciais estavam envolvidos. Ademais, seria o termo *remédio* algo mais popularmente conhecido, talvez mais eficiente para o *marketing* do autor. Algo similar ao que vemos hoje na TV e internet, nas publicidades atuais de uma indústria farmacêutica global. Propagandas que incentivam, no nosso país, a compra e o uso

¹ Fac-símile disponível no *site* do Museu da Farmácia de Portugal <https://www.museudafarmacia.pt/collection/ficha.aspx?lang=PO&a=1&pa=1&id=3562>.

de drogas sem prescrição médica, que ora são denominadas *remédios* ou *medicamentos*;

b) a exposição, por Semedo, de um valor de ciência para algo que deixaria de ser um simples *remédio*, *mistura*, *mezinha*, infusão, *alimento* ou *poção*. Com um valor agregado para tornar-se um *medicamento*. Conferiria-se, assim, em tese, mais prestígio ao autor como médico douto e um maior valor monetário ao produto designado como *medicamento*.

4.2. Graus de complexidade em contraste

Esperávamos que, havendo um verdadeiro (e concretizado) desejo de simplicidade textual, algo várias vezes expresso por Semedo, as métricas de complexidade dos textos do médico tenderiam a mostrar graus de complexidade apenas um pouco maiores em relação aos jornais escritos por pessoas e para comuns entre 1729-1731. Isto é, imaginamos que a escrita do médico, em 1707, já poderia concretizar uma tentativa de aproximação com um discurso popular escrito daquela época. Nessa linha de cotejos, tendo em mente que os parâmetros computacionais embutidos na ferramenta informatizada são os da atualidade, imaginamos que os textos de Semedo renderiam um bom contraste com textos jurídicos da atualidade. Esses dados apresentaremos mais adiante.

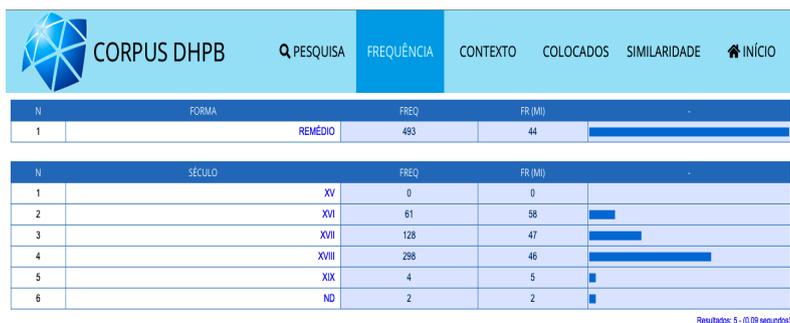
5. MEDICAMENTOS OU REMÉDIOS?

Conforme já apontou Gonçalves (2020, p. 70), ao estudar o primeiro manual de Enfermagem escrito e publicado em português, em 1741:

[...] «remédio» é, na *Arte de Enfermeiros*, o nome que recebe qualquer forma de tratar doenças ou aliviar os seus sintomas [...], indistintamente do tipo de substância ou produto (vegetal, animal, mineral ou outro) que entre na sua composição. [...] Por sua vez, a palavra «medicamento», que consta da nomenclatura lexicográfica de Bento Pereira (1697), segundo Houaiss (2001), já estaria atestada desde 1692.

Tendo consultado o *corpus* de base do DHPB (Biderman; Murakawa, 2021), observamos que os usos antigos de *medicamento* e *remédio* poderiam ser exemplarmente resumidos nas figuras 2 e 3 a seguir. Como se nota, a frequência média (FR(MI)) do uso do item *remédio* é significativamente **maior** (44) frente aos empregos de *medicamento* (06). Ainda assim, a maior concentração de ambos os itens dá-se em textos em português do século XVIII.

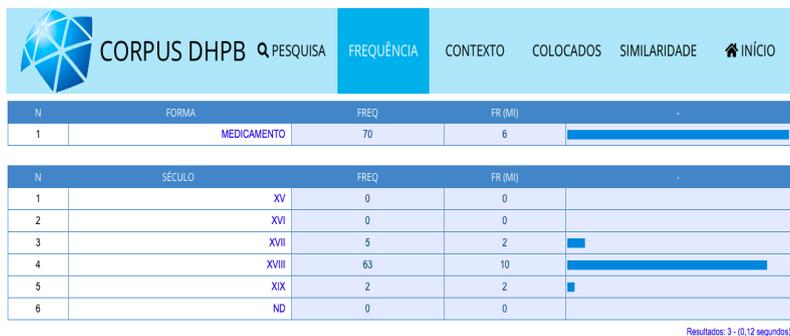
Figura 2 – Dados do item *remédio* no DHPB



VOLTAR PARA PESQUISA

Fonte: DHPB. Disponível em: <https://dicionarios.fclar.unesp.br/dhpb/>. Acesso em: ago. 2023.

Figura 3 – Dados do item *medicamento* no DHPB



Fonte: DHPB. Disponível em: <https://dicionarios.fclar.unesp.br/dhpb/>.
Acesso em: ago. 2023.

Vejamos agora, no Quadro 1, um resumo dos dados relacionados às 89 das 101 *Observações* médicas de Semedo, obtidos com a ferramenta AntConc (Anthony, 2014):

Quadro 1 – Dados de palavras mais frequentes em Semedo (1707) – até o item 20. Distribuição dos itens *remédio* (e variantes) e *medicamento* (e variantes), após os 20 primeiros itens mais frequente - destaque em negrito.

Semedo 1707 89 Observações – com AntConc #Word Types: 17.144 #Word Tokens: 146.905		
Ordem	Freq.	Item
1	7499	de
2	6379	que
3	4383	a
4	3528	o
5	2299	com
6	1932	os
7	1811	em
8	1597	as
9	1423	da

10	1409	naõ
11	1345	para
12	1338	do
13	1305	je
14	1300	por
15	1257	se
16	1129	porque
17	1084	ou
18	1072	como
19	899	huma
20	710	nem
32	539	remedios
37	456	remedio
940	18	remédios
1352	12	remédio
13575	1	muitosremedios
15157	1	remediorum
15883	1	talremedio
397	44	medicamentos
805	21	medicamento

Fonte: da autora

Em relação a uma proporção de usos ou de frequência dos itens em foco, somadas a suas formas variantes, considerando o total de palavras (146.905), temos o seguinte resultado: *remédio* (1.028 usos ou 0,069%), *medicamento* (65 usos ou 0,004%). Como se vê, a diferença segue sendo bastante significativa, com maciço emprego, por Seme-do, de ***remédio*** em vez de ***medicamento***.

A seguir, vale observarmos os registros do DHPB sobre as duas designações, com negritos nossos no interior dos verbetes, visando destacar os dois itens em estudo.

medicamento s.m.
variante: **medicam**

Remédio aplicável ao corpo, externa ou internamente.

Tomada a purga, que parecer mais conveniente, conforme for o fugeyto, e as forças opermittirem, fe tornar á a fazer outro fraſco de **medicamento** deobſtruyente na meſma fórma, que eſtá dito; [...]. LUIS GOMES FERREIRA (1735) [1735], [DASOBſTRUCÇOENS \[B00_0030 p. 63\]](#).

fig. Aquilo que diminui o ſoſrimento, os males da vida.

Porque como a abundancia e gula dos ricos é o ſeu veneno, e a eſtreiteza e abſtancia dos pobres o ſeu **medicamento**; os ricos chegam á ſepultura, como S.João á de Christo, [...]. PADRE ANTONIO VIEIRA (1951) [séc. XVII], [SERMÃO DA QUARTA DOMINGA DEPOIS DA PASCHOA \[A00_0908 p. 315\]](#).

1ª. datação [1618]

[...] ſômente com ſe lavar aquela parte três ou quatro vêzes com água morna , e quando ſe lhe não acode com eſte **medicamento** tão fácil, baſta aquela fogagem para vir a corromper todo o ſeſſo, com morte do enfermo, [...]. AMBRÓSIO FERNANDES BRANDÃO (1966) [1618], [DIÁLOGO SEGUNDO - QUE TRATA DO CLIMA E ENFERMIDADES DO BRASIL E DOS MEDICAMENTOS COM QUE SE CURAM \[A00_1582 p. 61\]](#).

remédio s.m.

variante: **remedio**.

Tudo que ſe toma ou ſe aplica para reparar ou reſtabelecer a ſaúde.

Caindo no princípio do mês paſſado de cama, e ſendo preciso no ſegundo tomar um **remédio** purgativo, chegou naquela meſma manhã a eſte arraial o Juiz de Fora que vai para o Mato Grosso [...]. FRANCISCO XAVIER DE

MENDONÇA FURTADO [1755], CARTA A SEBASTIÃO JOSÉ, TRATANDO DAS QUE RECEBERA DO BISPO E DOS OUVIDORES, DANDO-LHE NOTÍCIA DA REBELIÃO QUE O PADRE ROQUE INTENTARA FAZER NOS POVOS DA CIDADE DE BELÉM, COM AS SUAS PRÁTICAS ESPIRITUAIS(...) ARRAIAL DE MARIUÁ, 14 DE NOVEMBRO DE 1755 [A00_0367 p. 832].

fig. O que alivia e minora um mal moral.

De maneira que logo as molheres teriam **remedio** de vida, e estes homens remediarião suas almas, e facilmente se povoaria a terra. P. MANUEL DA NÓBREGA [1549], CARTA DO P. MANUEL DA NÓBREGA AO P. SIMÃO RODRIGUES, BAÍA 9 DE AGOSTO 1549 [A00_0002 p. 120].

1ª. datação [1548]

[...] pois della a ilha de Santa Helena nom pode aver mais de 300 leguoas que sam as paraijees em que por força as naos da Índia vão ter, como todalas cousas nos principios terão **remedio**, neste com tempo ho divia Vosa Alteza de por e deitar tão ma gente fora da terra e desta navegação ho que ao presente se pode fazer com pequeno custo e pouco trabalho, [...]. LUÍS DE GÓIS [1548], CARTA DE LUÍS DE GÓIS A D. JOÃO III, DANDO CONTA DOS PERIGOS QUE CORRIA O BRASIL COM A PRESENÇA DOS CORSÁRIOS FRANCESES, QUE JÁ CHEGAVAM À BAÍA DO RIO DE JANEIRO; VILA DE SANTOS, 12 DE MAIO DE 1548 (4) [A00_1588 p. 18].

Fonte: DHPB. Disponível em: <https://dicionarios.fclar.unesp.br/dhpb/>. Acesso em: ago. 2023.

Para ilustrar os usos do termo **remédio** (e suas variantes de escrita) em Semedo (1707), vejamos os seguintes excertos da sua *Observação I*. Esse é o primeiro relato do livro que contém o item, aqui reproduzido conforme um método particular que estabelecemos para

a transcrição do texto do volume. Quanto ao método de transcrição, ele pode ser considerado relativamente “peculiar”, pois foi ajustado aos nossos pontos de interesse, à necessidade de formatação para o processamento computacional e ao modo de trabalho colaborativo e didático, que envolveu diferentes turmas de estudantes. Cabe referir que: a) os números entre colchetes correspondem às linhas por página. Assim, a cada [1], indica-se o início de uma nova página do livro; b) os elementos marcados como colchetes angulados, tais como <par-tes>, correspondem à indicação da separação de sílabas conforme o original para quebra de linhas; c) os títulos das seções estão marcados com < >; d) sinais de pontuação são antecidos por um espaço; e) utilizam-se caracteres especiais, como o S longo em *Excellentiſſimo*; e) preservam-se as marcações originais de diacríticos e acentos, com em *hajaõ*; f) eventuais abreviaturas não foram estendidas ou desdobradas; g) citações em notas laterais em Latim estão assinaladas com < >, mas não foram transcritas. Os grifos em negrito servem apenas para guiar o olhar do nosso leitor.

<OBSERVAÇAMI .>

[1] De hũa Colica nephritica, que affligio ao Excellentiſſimo [2] Senhor Principe de Ligne, & Marquez de Arrouches <Arrou-ches> [3], em 6. de Janeiro de 1685 .

[...]

[31] Como pois eu tenho aſſentado que a verdadeira [32] ſciencia naõ consiste no eſtrondoso, ou campanudo das [33] palavras , nem na apparencia, ou pompa exterior dos [34] veſtidos , mas nas obras ordenadas com acerto, & eſſeituadas <eſſeituadas> [35] com felicidade , puz todo o cuidado em conhecer <co-nhecer> [36] a enfermidade, & os lugares em que eſtava , [37] para lhe applicar os **remedios convenientes** ; mas como [38] no

mesmo lugar (qual he o ventre) hajaõ muitas partes <par-tes>, [39] não he fácil distinguir o sitio da dor pela açcãõ leja <le-ja>,

[1] <ja>, nem pela retençãõ das fezes; porque tinha este Senhor [2] <Se-nhor> continuos defejos de vomitar, grandes dores no [3] ventre, excessiva adftriçãõ das fezes, repetidas vontades [4] de urinar, o que fazia gotta a gotta, & com grandiffimos [5] ardores. Porèm como estes symptomas sejaõ [6] quaõ identicos na dor de colica, & na nephritica, & [7] se equivoquem entre si de maneira, que a dor que começou [8] colica, acaba muitas vezes nephritica, (1.) <Citação de 9 linhas em latim do lado esquerdo> < 1. = nota lateral, ref. obras em latim> me [9] vi muito embaraçado para distinguir huma dor da outra [10]; porque sem ter perfeito conhecimento da doença <doen-ça> , [11] me exporia a algum ãinistro acontecimento: porque [12] se a dor fosse verdadeiramente colica, & eu lhe applicasse <ap-plicasse> [13] **remedios diureticos**, faria hum erro da primeira <pri-meira> [14] grandeza; porque levaria os humores para as [15] veas dos rins, que jaõ mais apertadas, & pequenas. (Semedo, 1707, p. 1-2).

Nessa mesma Observaçãõ, temos, mais adiante, a descriçãõ de um conjunto de *remédios* aplicados a pedido do paciente, destacados em vermelho:

[3] Neste tempo pedia o doente com enternecidos rogos <ro-gos>, [4] & suspiros lhe applicassem algum **remedio**, com [5] que se tiraassem as dores, ou se moderaasse tam cruel tormento <tor-mento>. [6] Para isso lhe mandei lançar huma ajuda de quatro <qua-tro> [7] onças de caldo de gallo velho, duas gemas de ovos [8] cruas, huma onça de terebinthina de betacom dez oitavas <oi-tavas> [9] de diaprunis; & sobre a parte dolorosa lhe appliquei

<apli-quei> [10] (alternadas vezes) hora um paõ vindo do forno [11] partido pelo meyo, & embebido em leite de cabras;

(...) [17] de erva fanta, & mangerona; porèm era taõ grande a pervicacia do mal, que nenhum alivio teve com algum [18] destes **remédios**, por mais anodynos, & decantados [19] que seirão (Semedo, 1707, p. 11-12).

Na sequência, na mesma *Observação I*, desta vez reproduzida sem marcas de transcrição, encontramos usos de **medicamento**, que parece ter alguma oposição com os *remédios* conforme antes apresentados. Conforme se nota, relaciona-se a **medicamento** ao preparo de uma **fórmula em pilulas**, compostas de substâncias específicas, e demais elementos destacados em negrito:

Aqui se esgotou o sofrimento do enfermo, & quasi desesperado se levantou da cama despido, & descomposto, andando de casa em casa com tam apressados movimentos, como se estivesse furioso, ou fóra de seu juizo. Neste lugubre espectáculo, & ultima desconfiança da vida o animei, dizendolhe que brevemente se veria restituindo à sua perfeita saude com hum **medicamento**, que em casos semelhantes me não deixára envergonhado. Ouvida esta alegre nova, renasceo no sobredito Senhor Marquez a esperança da vida, & pondo de parte toda a tristeza, & lembrança do muito que tinha padecido, tomou o **medicamento**, que **lhe mandei preparar na fórmula** seguinte. Tomem de calomelanos turqueti bem preparado meya oitava, de diagrydio fulfurado seis grãos, de laudano opiado, feito por mão de bom artifice, tres grãos; tudo se misture com o que bastar de confeição de hyacinthos, & formen-se **quatro pilulas**, que tem maravilhosa propriedade para render semelhantes dores, assim porque são narcoticas, pelo laudano

opiado que as fixa, & adormece; como porque são purgativas, pelo calomelanos, & diagrydio, com que são compostas. Tomadas **as ditas pilulas**, se tirou totalmente a dor, & dormio toda a noite com grande quietação, o que não podia fazer avia muito tempo & passadas seis horas acordou aos estímulos da purga, & obrou com tal felicidade, que ficou livre vinte & quatro horas, mas passadas estas, tornaraõ a vir as dores; donde vim a suspeitar que eu me tinha enganado no conhecimento da doença, como succedeo a Galeno (10.) (Semedo, 1707, p. 11-12).

Ainda assim, os modos de uso do termo *remédio* parecem alternar com um sentido de misturas ou preparados em geral, incluindo procedimentos como as *sangrias*. Mas, tal como parece ocorrer hoje, na linguagem cotidiana, não especializada, não ficava muito clara a diferença no uso contextualizado do século XVIII. No trecho a seguir, livre das marcas de nossa transcrição, Semedo apresenta três tipos ou gêneros de *remédios*. Vale sublinhar que, pelo que se lê, o autor considerava que as *pílulas* – antes referidas – feitas com as substâncias *calomelanos*, *tartaro* e *diagrydio* figuram também como *remédios*:

[...] não acharaõ que reprehender; antes se esprayaraõ em louvoires meus, de que sempre <me> confiderei indigno; & por isso de commum consentimento resolvè<r>aõ que usariaõ sómente de tres **generos de remedios**. O primeiro, feriaõ algumas sangrias nos pès, para que diminuida a carga dos humores, ficassẽ as vias mais patentes para expèllir melhor as areas, ou pedras, que presumiaõ aver nelle. O segundo, feria outra purga composta de calomelanos, tartaro, & diagrydio. O terceiro, feriaõ algũas amendoadas de pevides de melancia, & caroços de ginjas, & quinze grãos de semente de bijnaga feitas em agua cozida com folhas de pimpinella, nas quaes se desfariaõ dez grãos de

tartaro violado com huma onça dé lambedor de violas magiſtral; com os quaes **remédios** (repetidas [20] vezes tomados) lançou duas pedras, & algumas areas, & deſde aquelle dia deſapparecêraõ as dores, & ficou ſão (Semedo, 1707, p. 9).

Por fim, para um contraste do uso desses termos em textos antigos não especializados, ilustramos as primeiras ocorrências dos itens **remédio(s)** ao longo do volume 1 das citadas *Gazetas Manuscritas*. Essa forma, assim grafada, com acento e no plural, ocorre 7 vezes ao longo do volume. A forma no singular remédio ocorre 13 vezes. O item *medicamento*, sublinhamos, **não ocorre nesse volume 1 das Gazetas**, o que reforçaria a suspeita de *remédio* ser uma denominação popular. Todavia, percebe-se correlação com *caldas e/ou banhos terapêuticos*. Os números entre colchetes correspondem à linha do documento conforme transcrito na versão com ortografia modernizada, os grifos em negrito são nossos.

[89] **Diário de 20 de setembro de 1729** [111] Manuel de Oliveira da Cunha teve na Relação um acidente estando 24 horas sem acordo se restituiu com os remédios e vai as Caldas donde tem feito milagres as águas que Eliote descobriu a Senhora Condessa da Ericeira já está boa da queixa com que foi e volta brevemente e lá estão as Senhoras Dona Inácia de Rohan e a Senhora Dona Ana Joaquina. [100] Das Caldas há boas novas ainda que a Senhora Condessa da Ericeira vai devagar com a sua cura, e foi para tomar o mesmo **remédio** João Pedro Soares dando já cuidado as suas melancolias , e o acompanhou a Senhora Dona Ana Joaquina. [111] Manuel de Oliveira da Cunha teve na Relação um acidente estando 24 horas sem acordo se restituiu com os **remédios** e vai as Caldas donde tem feito milagres as águas que Eliote descobriu a Senhora

Condessa da Ericeira já está boa da queixa com que foi e volta brevemente e lá estão as Senhoras Dona Inácia de Rohan e a Senhora Dona Ana Joaquina.

Fonte: Gazetas Manuscritas. Disponível em: <https://books.openedition.org/cidehus/3083>.

Há também outros usos dignos de nota – em datas posteriores – sendo o segmento [1251] o que agrega *remédios* - em sentido figurado e literal e componto nome próprio, *pílulas* e o termo *sezões* (designação antiga para *febres*).

[313] Um capitão de mar e guerra se ofereceu com seus marinheiros a salvá-la, mas como se lhe não aceitou a proposta e o chamavam quando não tinha **remédio**, deu só o conselho de que metessem o navio a pique, porque seria fácil tirá-lo, salvando a artilharia, e apagando-se no mar o fogo, o enviado de Inglaterra passou um ofício justificando o referido .

[479] Queimaram-se algumas moradas de casas junto a **Nossa Senhora dos Remédios**; e pelos poucos oficiais que se acham livres, para acudir aos incêndios, se temeu que ardesse toda a rua mas supriu a boa diligência.

[819] O Senhor Infante Dom Francisco ainda está em Samora, e o Senhor Infante Dom Carlos esteve no último aperto, e foi por das vezes sangrado acrescentando-se o perigo com a invencível repugnância que tem a este, e outros **remédios** .

[1251] El-rei tomou **alguns remédios** em segredo e foram bem sucedidos a Senhora Infanta Dona Francisca está com um reumatismo o Senhor Infante Dom Carlos teve novos apertos e se lhe aumenta a debilidade o Senhor Infante Dom António veio de Pancas com **sezões** e já teve seis despediu-se-lhe **a febre** ontem tomou **pílulas** de

quina, e o Senhor Infante Dom Francisco o veio visitar e voltou logo da família de ambos adoeceram mais de 60 criados todos de **sezões** que dentro em Lisboa se tem diminuído muito.

Fonte: Gazetas Manuscritas. Disponível em: <https://books.openedition.org/cidehus/3083>.

6. ALGUMAS MEDIDAS SOBRE COMPLEXIDADE DO TEXTO MÉDICO ANTIGO

Dado que a ferramenta NILC-Metrix somente processa um texto por vez, destacamos aqui os resultados para a *Observação XLV*. Abaixo, temos seu trecho inicial, em itálico no original, também reproduzido sem nossas marcas de transcrição. Esse segmento inicial, presente em todas as 101 *Observações*, serve como um resumo do relato que seguirá em detalhes. Trazemos o excerto livre de nossas marcas e recursos de transcrição, conforme foi submetido a ambas as ferramentas computacionais. Destacamos, novamente em negrito, o uso de *remédio(s)*.

OBSERVAÇAM XLV.

De hum mercador, a quem repentinamente assaltou huma dor de colica taõ intoleravel, que estando na fé sacramental para commungar, o naõ pode fazer; e sendo eu chamado, conheci dos grandissimos ardores; e continuos desejos de ourinar, e vomitar, das picadas da bexiga, e do adormecimento da perna direita, que a tal dor era nephritica; para cujo **remedio** appliquei hum vomitorio de tres onças de agua benedicta vigorada , e tres ajudas feitas de cozimento de rim de vacca, misturando em cada huma, huma onça de terebinthina de beta lavada em agua de malvas, atè se fazer muito

branca, com huma gema de ovo crua; metendo depois disto ao doente em hum banho de agua, q̃ primeiro fosse cozida com hum arratel de amendoas doces bem pizadas; e foraõ estes **remedios** taõ maravilhosamente succedidos, que dentro de quatro horas deitou muitas pedras redondas do tamanho de grãos de pimenta, e no mesmo dia se tirãraõ as dores, e ficou saõ (Semedo, 1707, p. 275).

Conforme já mencionado, o sistema NILC-Metrix apresenta 200 medidas e/ou contagens sobre o vocabulário de um texto, que são agrupadas em 14 categorias. O foco principal dessas métricas é ajudar a calcular coesão, coerência e nível de complexidade textual, cujo escopo está detalhado em: <http://fw.nilc.icmc.usp.br:23380/metrixdoc>.

1. Medidas Descritivas
2. Simplicidade Textual
3. Coesão Referencial
4. Coesão Semântica
5. Medidas Psicolinguísticas
6. Diversidade Lexical
7. Conectivos
8. Léxico Temporal
9. Complexidade Sintática
10. Densidade de Padrões Sintáticos
11. Informações Morfossintáticas de Palavras
12. Informações Semânticas de Palavras
13. Frequência de Palavras
14. Índices de Leiturabilidade

Entre as 200 medidas, por questão de espaço, destacaremos aqui apenas as adiante resumidas, no Quadro 2. Comparam-se os valores gerados para a *Observação* 45 de Semedo com um segmento de semelhante número de palavras das *Gazetas Manuscritas*. Em vermelho, os

escores associados a uma, em tese, maior complexidade. Chamamos atenção para os valores da última coluna, denominada *Sentença Judicial*, que é um dos textos do tempo presente em cotejo. Essa Sentença integra todo um *corpus* sob exame no trabalho de doutoramento de Motta (2021) dedicado ao tema da complexidade dos textos jurídicos. Um tal contraponto com a atualidade pode parecer paradoxal, dado que temos um texto do século XXI, mas isso tem o propósito de destacar as condições do recurso computacional. A seguir, reproduzimos um segmento da Sentença para que se possa antever os elementos de complexidade envolvidos:

Passo a decidir.

INVERSÃO DO ÔNUS DA PROVA.

Inicialmente, considerando-se a controvérsia presente nos autos e, tendo em vista o disposto nos artigos 2º e 3º do CDC, tem-se que é aplicável a legislação consumerista ao caso em tela, dada a natureza da atividade que a requerida desenvolve, isto é, prestação de serviços de telefonia.

Nesse sentido, consigna-se, que há entendimento quanto à aplicação das disposições do Código de Defesa do Consumidor nas relações em que envolvam pessoa jurídica, com base na Teoria Finalista, já que se trata de destinatário final do serviço, sendo, portanto, parte vulnerável.

Desse modo, sendo aplicada as normas consumeiristas no caso em testilha, plenamente passível a inversão do ônus da prova, já que se trata de um direito básico do consumidor, visando impedir o desequilíbrio da relação jurídica, bem como a fim de facilitar a defesa desse no processo, uma vez que verificada a sua hipossuficiência.

Fonte: Motta, 2021, Anexos, *corpus* de estudo, disponível em: <https://github.com/search?q=Ester+-Motta&type=repositories>. Acesso em: 22 ago. 2023.

Quadro 2 – Resultados do sistema NILC-Metrix – em negrito, valores indicativos de tendência de **maior** complexidade entre os dois textos antigos, Semedo e Gazetas, e o texto moderno

Categoria da métrica	Nome da métrica	Descrição da métrica	Escore Semedo	Escore Gazetas	Sentença Judicial atual
Simplicidade Textual	dialog_pronoun_ratio	Proporção de pronomes pessoais que indicam uma conversa com o leitor em relação à quantidade de pronomes pessoais do texto Interpretação: quanto maior a métrica, menor a complexidade	1.0	0	0
Simplicidade Textual	easy_conjunctions_ratio	Proporção de conjunções fáceis em relação à quantidade de palavras do texto Interpretação: quanto maior a métrica, menor a complexidade	0.141	0.097	0.067
Simplicidade Textual	hard_conjunctions_ratio	Proporção de conjunções difíceis em relação à quantidade de palavras do texto Interpretação: quanto maior a métrica, maior a complexidade	0.022	0.017	0.037

Categoria da métrica	Nome da métrica	Descrição da métrica	Escore Semedo	Escore Gazetas	Sentença Judicial atual
Simplicidade Textual	long_sentence_ratio	Proporção de Sentenças Muito Longas em relação a todas as sentenças do texto Interpretação: quanto maior a métrica, maior a complexidade	0.489	0.579	0.613
Índices de Leiturabilidade	dalehall_adapted	Fórmula Dale Chall adaptada Interpretação: quanto maior o valor da métrica, maior a complexidade textual	13.823	12.068	12.364
Índices de Leiturabilidade	flesch	Índice Flesch Interpretação: quanto maior o resultado da métrica, menor a complexidade textual.	41.039	52.997	33.091
Índices de Leiturabilidade	gunning_fox	Índice Gunning Fog Interpretação: quanto maior a métrica, maior a complexidade	17.395	10.541	9.297
Índices de Leiturabilidade	honore	Estatística de Honoré Interpretação: quanto mais alto o valor , mais rico o texto é lexicalmente, o que está associado com maior complexidade .	1035.354	1036.178	967.730

Fonte: elaboração da autora

Os resultados obtidos indicam que o texto de Semedo tenderia a ser qualificado como **MENOS** complexo do que o texto do jornal manuscrito antigo quanto às seguintes medidas:

- Proporção de pronomes pessoais que indicam uma conversa com o leitor em relação à quantidade de pronomes pessoais do texto (quanto maior o valor, mais complexo)

Interpretação da métrica no sistema: textos que estabelecem diálogo com o leitor têm menor complexidade, portanto, quanto maior o resultado, menor a complexidade.

Descrição da métrica: proporção de pronomes pessoais que indicam uma conversa com o leitor em relação ao total de pronomes pessoais presentes no texto.

Definição dos termos que aparecem na descrição da métrica: pronomes pessoais que indicam uma conversa com o leitor são: “eu”, “tu”, “você” e “vocês”.

Exemplo (Semedo, 1707, p. 11, Obs. I):

Fallo verdade, creame quem quizer, que **eu** não obrigo a alguem a que me crea.
Se com tudo ouver algum Medico tam pertinaz, & inflexivel, que reprove este modo de curar, veja as Observações de Riverio [...]

Por outro lado, o texto de Semedo seria **MAIS** complexo do que o texto das *Gazetas* quanto a:

- Proporção de conjunções difíceis em relação à quantidade de palavras do texto (quanto maior o valor, mais complexo)

Descrição da métrica no sistema: as conjunções difíceis do sistema são aquelas que estão em um conjunto de referência. Esse conjunto compreende os elementos: todavia, eis, a fim de, ao passo que, para que, conforme, tais, ou seja, contudo, bem como, logo, à medida que, entretanto, desde que, mesmo que, ainda que, de acordo com, uma vez que, por sua vez, sobretudo, até, ainda, caso, no entanto, nem, quanto, já, como, já que, outrossim, mas também, como também, não só, mas ainda, tampouco, senão também, bem assim, ademais, antes, não obstante, sem embargo, ao passo que, de outra forma, em todo caso, aliás, de outro modo, por conseguinte, em consequência de, por consequência, consequentemente, conseguintemente, isso posto, pelo que, de modo que, de maneira que, de forma que, em vista disso, por onde, porquanto, posto que, isto é, ademais, senão, dado que, visto como, vez que, de vez que, pois que, agora, na medida em que, sendo que, como que, como quer que, eis que, sendo assim, tal qual, ao invés de, conquanto, por muito que, visto que, uma vez que, quanto mais, quanto menos, se bem que, apesar de que, suposto que, ainda quando, quando mesmo, a despeito de, conquanto que, sem embargo de que, por outro lado, em contrapartida, sem embargo, muito embora, inclusive se, por mais que, por menos que, por pouco que, contanto que, salvo se, com tal que, caso que, consoante, tal que, de forma que, à proporção que, ao passo que, mal, tão logo, entretanto, sob esse aspecto, sob esse prisma, sob esse ponto de vista, sob esse enfoque, embora, portanto, além disso.

Exemplo (Semedo, 1707, p. 11, Obs. II):

[...] o Medico, que não se compadece do seu proximo estando enfermo, he indigno de ter nem o nome, nem a apparencia de homem; & sem **embargo** que para empresas grandes não basteão forças pequenas, & eu conheço as poucas que tenho, nem por isso me escusei já mais de fazer

o que entendo que he conveniente aos doentes ,
antepondo as Juas vidas ao meu credito.

- Fórmula Dale Chall adaptada

Descrição da métrica no sistema: combina a quantidade de palavras não familiares com a quantidade média de palavras por sentença. As “palavras não familiares” são as que não constam do vocabulário básico conhecido por alunos do quarto ano do Ensino Fundamental do Brasil. Foram utilizadas como referência de vocabulário familiar ou acessível as entradas de um *Dicionário de Palavras Simples* de Maria T. C. Biderman (2006).

- Índice Flesch (quanto **menor** o valor, mais complexo)
- o Índice de Leiturabilidade de Flesch busca uma correlação entre tamanhos médios de palavras e sentenças.
- Índice Gunning Fog

Descrição da métrica: o índice de leiturabilidade Gunning Fog (também conhecido como Gunning FoX) soma a quantidade média de palavras por sentença ao percentual de palavras difíceis no texto e multiplica tudo por 4. O resultado está diretamente ligado aos 12 níveis do ensino americano. Índices superiores a 12 representam textos extremamente complexos.

Definição dos termos que aparecem na descrição da métrica no sistema: palavras difíceis, para essa métrica, são aquelas que possuem mais de duas sílabas.

Limitações da métrica: nem sempre as palavras com mais de duas sílabas são difíceis em português (ex: árvore, professor, escola, salada).

Para as *Gazetas*, quando revelam uma complexidade que é **MAIOR** do que do texto de Semedo, temos as seguintes medidas distintivas:

- Proporção de conjunções fáceis em relação à quantidade de palavras do texto – cuja lógica é semelhante à do conjunto de itens difíceis.

Como palavras e expressões fáceis, temos: como, se, mas, quando, ou, que, porque, e, assim, porém, caso, por isso que por isso, por enquanto, enquanto isso, enquanto, pois, além de, então, daí, por exemplo, ou seja, sem que, para que, cada vez que, antes que, assim como, tanto quanto, feito, que nem, toda vez que, a não ser que, depois que, até que, na medida em que, desde, nem bem, tanto que, segundo, assim que, tanto que, tão que, sem que, ora.

- Proporção de sentenças muito longas em relação a todas as sentenças do texto
- Estatística de Honoré

Descrição desta métrica no sistema: a estatística de Honoré é um tipo de *type/token ratio* – medida de diversidade do vocabulário baseada na proporção de formas únicas em relação ao total de palavras do texto – mas que leva em consideração, além da quantidade de *types* e *tokens*, a quantidade de *hapax legomena*.

Em síntese, observamos que, como o sistema baseia-se nas normas e acervos do português atual, isso deve ser ponderado, sobretudo porque lidamos com diferentes padrões de ortografia antiga e transcrições, mais e menos modernizadas. A medida Dale Chall adaptada para o português (métrica de leiturabilidade, número 196),

por exemplo, combina a quantidade de palavras desconhecidas com a média de palavras por frase. As “palavras desconhecidas” são aquelas que *hoje* não fazem parte do vocabulário básico conhecido pelos alunos do quarto ano do Ensino Fundamental brasileiro. Limitação semelhante pode ser representada pela medida *Gunning Fog Readability Index* (199), afinal as “palavras difíceis” são aquelas que possuem mais de duas sílabas, o que seria um padrão da língua inglesa e não do português.

Por outro lado, há medidas que se poderiam considerar independentes do tempo ou do período de produção dos textos sob exame, como o Índice de *Leiturabilidade de Flesch*. A medida *Honoré* seria outro exemplo, pois é um tipo de relação *type/token* considera, além da relação de palavras diferentes/total de palavras, também a quantidade de *hapax legomena*, que são as palavras de única ocorrência ao longo do texto. Esse é o conjunto de palavras que tende a conformar, em termos estatísticos, a maior dimensão da variabilidade do vocabulário de um texto. Mas é preciso não esquecer que lidamos com diferentes grafias e formas das palavras nos textos antigos, o que incrementa o número de *types* e as medidas de variabilidade lexical baseadas em formas, como em remédio com e sem acento ou **agua** e **agoa**, que podem ocorrer em um mesmo texto-fonte.

Enfim, há muito o que se ponderar com relação a um universo dessas 200 métricas que um sistema atual nos oferece a partir de um texto antigo e mesmo de um texto atual como uma Sentença Judicial. Nesse caso, esse tipo textual e discursivo atual serviu para nos mostrar uma complexidade global-média, em tese, mais elevada do que as dos textos antigos de Semedo, especialmente pelas medidas Índice de *Leiturabilidade de Flesch*; *Maior presença de conjunções difíceis* e *proporção de sentenças muito longas*. Esse texto atual também serve para ilustrar o

tamanho do nosso “improviso” ao recorrer a um sistema atual - abastecido com dados do português atual - para lidar com textos antigos.

Ampliando os contrapontos, por outro lado, imaginamos que seria relevante explorar traços de um distanciamento entre a escrita de Semedo e textos tidos como moderadamente complexos naquela época, como, por exemplo, seriam os textos dos Sermões do Pe. Antonio Vieira, produzidos entre 1679-1695. Tendo iniciado esse contraste Semedo-Vieira, com a mesma ferramenta NILC-Metrix, teste que não será aqui detalhado, já identificamos indícios de uma grande proximidade, em termos globais, entre ambos². Isto é, *grosso modo*, podemos pensar que a escrita do médico alentejano ainda se mantivesse bastante presa a um estilo barroco que o precede, embora Semedo, como outros da sua época, reconhecesse a necessidade de acessibilidade. Não obstante, nesse contraponto inicial, percebemos tendência à maior facilitação do que Vieira em diferentes aspectos da tessitura textual, como a sintaxe, por exemplo.

CONSIDERAÇÕES FINAIS

Conforme pudemos demonstrar, *medicamento* e *remédio* são formas que eram empregadas de modos bastante semelhantes nos nossos documentos antigos, fossem eles textos especializados ou não. Um viés de significação farmacológica mostrou-se bastante discreto em diferentes usos nos materiais antigos. Assim, podemos imaginar “nichos de significação” que se entrecruzam também no passado. Esses nichos podem ser conectados à ideia atual de uma teoria de *frames semânticos*. Todavia, na contemporaneidade dos empregos

² Esse estudo comparativo entre obras do Pe. Antonio Vieira e outras obras médicas do século XVIII abastece um artigo que estamos elaborando com colegas pesquisadores de PLN.

dessa dupla terminológica, ainda que se procure por uma diferença conceitual e científica pontual entre ambos, alternâncias e sobreposições persistem. Desde o século XVIII, **remédio** mostra-se como uma designação mais popular, forma à qual se apela, hoje, principalmente na publicidade, quando é conveniente ou necessário um entendimento mais facilitado da informação, em que pese a distinção ser necessária. O item **medicamento** é mais recente e tem um caráter mais erudito ou ‘técnico’, fator que a Farmacologia atual sublinha e salvaguarda. Em termos científicos, bem sabemos, **remédio** não é o mesmo que **medicamento**. Ainda assim, a forma **remédio** parece ser a mais acessível, em função do fator frequência ser maior nos dias de hoje, conforme o percurso histórico já mostrava. Pesquisando no CorPop³, um *corpus* de referência do português popular escrito atual, encontramos que **remédio(s)** tem 59 ocorrências contra apenas 21 de **medicamento(s)**. Nesse *corpus*, os usos do item **remédio(s)** parecem ser bastante amplos, englobando a especificidade técnica que deveria conectar-se à outra denominação, como nesse exemplo: “Em laboratório, é usado em **remédios** que combatem inflamação e alergia”.

Quanto a uma apreensão de traços de provável complexidade textual – em textos antigos, via medidas de uma moderna ferramenta computacional, o texto de Semedo cumpre a suposição de que tentaria ser, em tese, relativamente próximo das *Gazetas*. Todavia, mantinha-se um tanto (bem) mais complexo do que elas. O texto do médico seria mais difícil do que aqueles jornais de sua época quando observamos as medidas de *Leiturabilidade de Flesch* (quanto maior o valor, menos complexo) e o Índice Gunning Fog (quanto maior, mais complexo).

³ Cf. *Corpus* de acesso gratuito em: <http://www.ufrgs.br/textecc/porlexbras/corpop/index.php>, veja a opção Caixa de Ferramentas.

Do mesmo modo, a proporção de sentenças muito longas em relação a todas as sentenças do texto do médico tende a ser relativamente próxima à das *Gazetas*. Os elementos associados à presença de pronomes marcadores de diálogo e de conjunções difíceis precisariam ser relativizados em meio às convencionalidades discursivas naquela época. Assim, em síntese, o texto do médico parece ser, sim, tentativamente facilitado, pois aproximava-se de parâmetros de facilitação que imaginamos para as *Gazetas*. Entretanto, há pontos de distanciamento bem marcados e pontuais, mostrando-se, em tese, o texto médico mais complexo ou mais difícil do que os jornais seriam. Situação que fica marcada pelo Índice Flesch, que seria uma medida de facilidade do texto relativamente independente de tempo.

Para finalizar este capítulo, queremos comentar sobre alternativas para maior fomento de estudos sobre a história das linguagens especializadas em seus diferentes aspectos, do léxico à sua ATT, incluindo especificidades sintáticas. A presença de uma interface computacional - e de outras tantas interfaces - envolvidas nesse tipo de estudo linguístico e lexicológico - mostram traços da sua “modernidade” e as tantas interdisciplinaridades envolvidas e necessárias. Além disso, deve ficar claro que há toda uma contraparte tecnológica para apoiar o insubstituível trabalho humano do analista da linguagem.

O processamento do português, como um todo, seja o atual ou o mais antigo, é um dos desafios da computação em grande escala do século XXI. Estudos em torno de textos médicos antigos, do ponto de vista histórico, técnico-científico, socioeconômico, de difusão e de inovação tecnológica, mostram outros tantos novos desafios para o moderno processamento do português e de suas ferramentas. Esses desafios, conforme registrado no livro *A Língua Portuguesa na Era Digital* (Branco *et al.*, 2012), também incluem sabermos lidar com acervos documentais do passado, em seus suportes, grafias e valores,

em tese, tão diferentes dos atuais. Além disso, vale reforçar, é “*da maior importância a investigação realizada na área da tecnologia da linguagem para a consolidação do português como língua de comunicação global na sociedade da informação*” (cf. Dr. Pedro Passos Coelho, Primeiro-Ministro de Portugal, citação da contracapa de Branco *et al.*, 2012). E aqui temos as heranças e todo um patrimônio histórico legados em português, que não podem ser desprezados.

A descrição de conteúdo lexical e terminológico de textos científicos e pré-científicos antigos, como os textos médicos do século XVIII, relacionados aos temas de doenças e seus tratamentos, tende a revelar pontos de conhecimento ainda válidos: conhecimentos históricos, culturais, químicos e farmacológicos, por exemplo. Saberes que se construíram no cenário de uma lusofonia que se afirmava. Em meio aos revezes do século XXI, é um alento poder apreciar, em panorama, os empreendimentos de um século passado cujos ideais buscavam a disseminação e a facilitação de conhecimentos e o equacionamento entre saberes populares, científicos e dogmas. Assim, esses passados também nos ensinam caminhos para o futuro, além de mostrarem o quanto *remédios* também eram *medicamentos*.

REFERÊNCIAS

ANTHONY, Lawrence. **AntConc (Version 3.4.3) [Software]**. Tokyo: Waseda University, 2014. Disponível em: <http://www.lauren-ceanthy.net/>. Acesso em: 23 ago. 2023.

BARBOSA, Ana V. B. Do conhecimento da doença à sua nomeação: uma viagem pelo Tratado da Conservação da Saude dos Povos, de António Ribeiro Sanches. **Panace@**, Córdoba, v. 21, n. 52, p. 37-48, jul./dez. 2020. Disponível em: https://www.tremedica.org/wp-content/uploads/panacea20-52_08_Tribuna_VieiraBarbosa.pdf. Acesso em: 23 ago. 2023.

BERBER SARDINHA, Tony. **Linguística de Corpus**. Barueri: Manole, 2004.

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

BIDERMAN, Maria Tereza C.; MURAKAWA, Clotilde de Almeida Azevedo (Orgs.). **Dicionário Histórico do Português do Brasil**: séculos XVI, XVII e XVIII. Araraquara: FCL-UNESP, 2021. Disponível em: <http://dicionarios.fclar.unesp.br>. Acesso em: 15 out. 2021.

BRANCO, António Branco; MENDES, Amália; PEREIRA, Sílvia; HENRIQUES, Paulo; PELLEGRINI, Thomas; MEINEDO, Hugo; TRANCOSO, Isabel; QUARESMA, Paulo; LIMA, Vera Lúcia Strube de; BACELAR, Fernanda. **A Língua Portuguesa na Era Digital / The Portuguese Language in the Digital Age**. White Paper Series. Berlim: Springer, 2012. Disponível em: <http://metanet4u.eu/wbooks/portuguese.pdf>. Acesso em: 28 ago. 2023

CAMPOS, Rafael Dias da Silva. Saúde em Portugal na virada dos séculos XVIII e XIX, o caso dos Avisos ao Público. *In*: POLÓNIA, Amélia; BRACHT, Fabiano; CONCEIÇÃO, Gisele Cristina da; PALMA, Monique. **História e Ciência**: Ciência e Poder na primeira Idade Global. Porto: Faculdade de Letras da Universidade do Porto, 2016. p. 40-57. Disponível em: <https://repositorio-aberto.up.pt/bitstream/10216/87047/2/163483.pdf>. Acesso em: 23 ago. 2023.

CORPUS **Dicionário Histórico do Português do Brasil** (DHPB). Araraquara: FCL-UNESP, 2021. Disponível em: <http://dicionarios.fclar.unesp.br>. Acesso em: 23 ago. 2023.

DURY, Pascaline; PICTON, Aurélie. Terminologie et diachronie: vers une réconciliation théorique et méthodo-logique? **Revue française de linguistique appliquée**, Paris, v. 14, p. 31-41, 2009.

FINATTO, Maria José B. **Definição terminológica**: fundamentos teórico-metodológicos para sua descrição e explicação. 2001. 395f. Tese (Doutorado em Estudos da Linguagem, Teorias do Texto e do Discurso) – Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2001.

FINATTO, Maria José B. Medicina em português no século XVIII: desafios da Terminologia Diacrônica no cenário das Humanidades Digitais. **Panacea**, Córdoba, v. 21, n. 52, p. 20-36, jul./dez. 2020. Disponível em: https://www.tremedica.org/wp-content/uploads/panacea20-52_07_Tribuna_BocornyFinatto.pdf. Acesso em: 23 ago. 2023.

FINATTO, Maria José B.; QUARESMA, Paulo; GONÇALVES, Maria Filomena. Portuguese *corpora* of the 18th century: old medicine texts for teaching and research. *In*: FISER, Darja; PANCUR, Andrej (ed.). **Proceedings**

of the Conference on Language Technologies and Digital Humanities, Faculty of Electrical Engineering, University of Ljubljana: Ljubljana, Slovenia, Sep. 20-21, 2018. p. 114-120. Disponível em: http://www.sdjt.si/wp/wp-content/uploads/2018/09/JTDH-2018-Jose-et_al_Portuguese-Corpora-of-the-18th-century-old-Medicine-texts-for-teaching-and-research-activities.pdf. Acesso em: 23 ago. 2023.

GAZZOLA, Murilo Gleyson; LEAL, Sidney Evaldo; ALUÍSIO, Sandra Maria. Predição da complexidade textual de recursos educacionais abertos em português. In: **Anais do Symposium in Information and Human Language Technology - STIL**. Porto Alegre: SBC, 2019. Disponível em: <http://comissoes.sbc.org.br/ce-pln/stil2019/proceedings-stil-2019-Final.pdf>. Acesso em: 23 ago. 2023.

GONÇALVES, Maria Filomena. A Arte de Enfermeiros (1741): aspectos do léxico relativo a doenças e remédios no século XVIII. **Panacea@**, Córdoba, v. 21, n. 52, p. 68-85, jul./dez. 2020. Disponível em: <https://www.tremedica.org/panacea/6278/>. Acesso em : 23 ago. 2023.

GOLUB, Koraljka; LIU, Ying-Hsang. **Information and Knowledge Organisation in Digital Humanities: Global Perspectives**. 1. ed. Oxfordshire: Routledge, 2021. Disponível em: <https://doi.org/10.4324/9781003131816>. Acesso em: 23 ago. 2023.

JASMIN, Marcelo Gantus. História dos conceitos e teoria política e social: referências preliminares. **Revista Brasileira de Ciências Sociais**, São Paulo, v. 57, n. 20, p. 27-38, 2005.

KABATEK, Johannes. sobre a historicidade de textos. Trad. por: José da Silva Simões. **Linha d'água**, São Paulo, v. 17, p. 157-170, 2005.

LISBOA, João L.; MIRANDA, Tiago. C. P. dos Reis; OLIVAL, Fernanda. **Gazetas manuscritas da Biblioteca Pública de Évora**. v. I (1729-1731) Lisboa: Edições Colibri, CHC-UNL e CIDEHUS-EU, 2002/2018 (on-line). Disponível em: <https://books.openedition.org/cidehus/3083>. Acesso em: 23 ago. 2023.

LOURENÇO, Tânia Souza. **O médico entre a tradição e a inovação**: João Curvo Semedo. 2016. Dissertação (Mestrado em História) – Universidade Federal Fluminense, Niterói, 2016.

MARENGO, Sandro Marcio Drumond Alves. **Variações terminológicas e diacronia**: estudo léxico-social de documentos manuscritos militares dos séculos XVIII e XIX. 2016. Tese (Doutorado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Belo Horizonte, 2016.

MAURAN, M. G. **Aviso à gente do mar sobre a sua saúde**. Traduzido por B. J. Carvalho. Lisboa: Typ. de João Antonio da Silva, 1794.

MOTTA, Ester. Índices de Complexidade Textual em Sentenças dos Juizados Especiais Cíveis do Poder Judiciário do Estado do Rio Grande do Sul. **Inventário**, Salvador, v. 1, p. 35-50, 2018. Disponível em: <https://periodicos.ufba.br/index.php/inventario/article/view/23570> Acesso em: 25 ago. 2023.

MOTTA, Ester. Sentenças Judiciais e Acessibilidade Textual e Terminológica. **Domínios de Linguagem**, Uberlândia, v. 15, n. 3, p. 761-813, ago. 2021. Disponível em: <https://seer.ufu.br/index.php/dominiosdelinguagem/article/view/52909> Acesso em: 25 ago. 2023.

MURAKAWA, Clotilde de Almeida Azevedo. Vocabulário das enfermidades em documento do Brasil colonial: o relato de “Prodigiosa Lagoa” (1749). In: MURAKAWA, Clotilde de Azevedo Almeida; NADIN, Odair Luiz Nadin (orgs.) **Terminologia: uma ciência interdisciplinar**. São Paulo: Cultura Acadêmica, 2014. p. 83-101.

NILC-METRIX. Disponível em: <http://fw.nilc.icmc.usp.br:23380/nilcmetrix> Acesso em: 20 ago. 2023.

PALMA, Monique. Textos e contextos sobre cirurgia em Portugal e na América portuguesa no século XVIII. **Magallánica: revista de historia moderna**, Mar del Plata, v. 7, n. 14, p. 335-361, 2021. Disponível em: <https://fh.mdp.edu.ar/revistas/index.php/magallanica/article/view/5286>. Acesso em: 18 ago. 2023.

PAIXÃO DE SOUSA, Maria Clara; TRIPPEL, Thorsten. Building a historical corpus for Classical Portuguese: some technological aspects. In: INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION (LREC), 5., 2006. **Proceedings** [...]. Genova: ELRA, 2006. p. 1831-1836.

QUARESMA, Paulo; FINATTO, Maria José B. Information Extraction from Historical Texts: a Case Study. In: DHNLP - PROPOR. **Workshop on Digital Humanities and Natural Language Processing**. Évora – Portugal: CEUR ORG., 2020. Disponível em: <http://ceur-ws.org/Vol-2607/short2.pdf>. Acesso em: 25 ago. 2023.

QUARESMA, Paulo; FINATTO, Maria José B.; VIEIRA, Renata; TROJAHN, Cássia. **Extração de informação em manuais de medicina do século XVIII: estudo de caso com dados brutos**. Évora: Universidade de Évora -

CIDEHUS, no prelo (2022). Disponível (em breve) em: <https://books.openedition.org/cidehus/>. Acesso em: 20 ago. 2023.

SANTOS, Ivo; OLIVAL, Fernanda; SEQUEIRA, Ofélia. Excavating the Data Pit: the Portuguese Parish Memories (1758) as a Gold Standard. *In: DHNLP - PROPOR. **Workshop on Digital Humanities and Natural Language Processing***. Évora – Portugal: CEUR ORG., 2020. p. 69-75. Disponível em: <https://dblp.org/db/conf/dhandnlp-ws/dhandnlp-ws2020.html>. Acesso em: 20 ago. 2023.

SEMEDO, João Curvo. **Observações medicas doutrinaes de cem casos gravissimos, que em serviço da pátria, & das nações estranhas escreve em lingua portugueza, & latina**. Lisboa: Officina de Antonio Pedrozo Galram, 1707. 635p.

TEMMERMAN, Rita. Une théorie réaliste de la terminologie: le sociocognitivisme. **Terminologies nouvelles**, Belgique, v. 21, p. 58-64, 2000.

O SANGUE É UM LÍQUIDO OU UM TECIDO? A ELABORAÇÃO DE DEFINIÇÕES PARA UM DICIONÁRIO TERMINOLÓGICO DAS CIÊNCIAS NATURAIS PARA PROFESSORES

*Mariângela de Araújo
Ariane Vicente Mota*

INTRODUÇÃO

As Ciências Naturais configuram-se como uma grande área do saber que reúne, entre outras áreas, a Biologia, a Física e a Química. Essa grande área, além de ter grande tradição histórica, uma vez que a vida, o ambiente e o universo sempre suscitaram na humanidade grandes questionamentos, atualmente mantém o seu vigor, contribuindo grandemente para o entendimento do mundo em que vivemos e seus elementos e para o desenvolvimento científico e tecnológico. Além disso, apresenta à sociedade uma série de estudos e reflexões essenciais às decisões e ao comportamento do homem em relação à natureza.

Nesse sentido, é possível compreender que seja relevante que esses questionamentos e reflexões estejam presentes desde os primeiros anos do ensino fundamental, uma vez que as crianças, além de muito curiosas a respeito do mundo que as cerca, devem ser levadas

a refletir sobre questões essenciais à preservação da natureza e à manutenção da vida em nosso planeta. Não se pode também esquecer que a formação de futuros estudiosos e pesquisadores das Ciências Naturais inicia-se nesse nível de ensino, o que faz ser primordial que o processo de ensino ocorra de forma adequada, por meio de uma abordagem correta dos conceitos que irão acompanhar essas crianças em seu processo educacional.

Apesar dessa relevância, ainda se assiste, no Brasil, a certa negligência com a formação dos professores que iniciarão o ensino desses estudantes, sobretudo no que diz respeito às Ciências Naturais.

Tal fato fica evidente ao se verificar que na versão atualizada da LDB, em 2020, no que diz respeito à formação dos professores do primeiro ciclo do ensino fundamental, ainda se mantenha, da mesma maneira que na versão anterior (1996), como exigência mínima de formação o Ensino Médio:

Art. 62. A formação de docentes para atuar na educação básica far-se-á em nível superior, em curso de licenciatura plena, admitida, como formação mínima para o exercício do magistério na educação infantil e nos cinco primeiros anos do ensino fundamental, a oferecida em nível médio, na modalidade normal (LDB, 2020).

Em estudo anterior desenvolvido por Araújo e Souza (2011), verificou-se que, apesar de haver essa exigência pela LDB de 1996, no *Plano Nacional da Educação* publicado em 2001, no ano de 1996, aproximadamente noventa e cinco mil (95.000) funções docentes eram exercidas por professores que não apresentavam essa formação mínima. Ainda nesse documento podia-se observar que, de um total de setecentos e noventa e oito mil, novecentas e quarenta e sete (798.947) funções docentes, apenas cento e setenta e dois mil sete-

centas e quinze (172.715) eram exercidas por professores que tinham formação em nível superior, ou seja, pouco mais de vinte por cento. Além disso, dentre as metas e diretrizes apontadas, o documento indica o prazo de cinco anos para a formação mínima exigida de todos os professores, ou seja, ensino médio na modalidade *Normal*, e um prazo de dez anos para a formação de 70% dos professores em nível superior específico. Partindo, então, da data do documento, esperava-se que em 2011 todos os professores desse nível de ensino estivessem formados em nível superior.

Entretanto, ainda no referido estudo, Araújo e Souza mencionam a Portaria 10, de 06 de agosto de 2009, do Conselho Nacional de Educação (CNE), intitulada *Indicações para subsidiar a construção do Plano Nacional de Educação 2011-2020*, que indica, como uma das prioridades na formação dos profissionais, “Estabelecer um prazo para extinguir o curso normal de nível médio no País, para que ele deixe de ser considerado como formação inicial do professor e, conseqüentemente, como patamar básico de remuneração” (p. 37). Tal afirmação pressupõe que, em 2009, a dois anos do prazo estabelecido pelo PNLD de 2001, a questão não estava superada e o prazo para se atingir a meta deveria ser revisto.

Conforme previsto, no *Plano Nacional de Educação* mais recente, aprovado pela lei nº 13.005/2014, com vigência de 10 anos, novamente encontra-se uma meta relacionada à formação de professores:

META 15 Garantir, em regime de colaboração entre a União, os Estados, o Distrito Federal e os Municípios, no prazo de 1 (um) ano de vigência deste PNE, política nacional de formação dos profissionais da educação de que tratam os incisos I, II e III do caput do art. 61 da Lei nº 9.394, de 20 de dezembro de 1996, assegurado que todos os professores e as professoras da educação básica possuam

formação específica de nível superior, obtida em curso de licenciatura na área de conhecimento em que atuam.

No *Relatório do Primeiro Ciclo de Monitoramento das Metas do PNE: Biênio 2014-2016* (p. 358) – o mais recente disponibilizado –, lê-se que:

O Indicador 15 aponta o percentual de docências com professores que possuem formação superior compatível com a disciplina que lecionam. No ano de 2015, esse percentual foi de 52,5% para a educação básica. Considerando cada uma das etapas, a maior compatibilidade foi observada no ensino médio, cujo resultado para esse indicador foi de 56,6%. Em seguida, tivemos os anos iniciais do EF, com 55,7%. Ficaram abaixo de 50,0% de compatibilidade os anos finais do EF (47,0%) e a educação infantil (44,6%).

Esses últimos documentos demonstram que o Brasil ainda não cumpriu sua meta de ter professores com formação em nível superior compatível com as disciplinas que lecionam. Isso vale para todo o ensino básico. Todavia, é possível perceber nas atuais descrições dessa situação certa falta de transparência em relação aos dados, uma vez que, tanto nas metas divulgadas no PNE de 2014 quanto no Relatório referente a 2014-2016, não é colocado claramente o nível de formação dos professores, ou seja, não é possível saber se a “falta de compatibilidade” ocorre devido ao nível de formação ou à área de formação em nível superior. De toda forma, é possível verificar que mais de 44% das docências nas séries iniciais do ensino fundamental ainda são ministradas por professores que não apresentam formação superior em Pedagogia – considerada pelos documentos como a ideal para esse nível.

Voltando ao estudo de Araújo e Souza (2011), segundo seus levantamentos, é possível ainda verificar que não há no âmbito do documento oficial que regulamenta o oferecimento de cursos superiores para a formação de professores do ensino fundamental nas séries iniciais (*Resolução CNE/CP 1*, de 18 de fevereiro de 2002), menção sobre a carga horária mínima destinada aos conteúdos referentes às Ciências Naturais.

Desse modo, verifica-se que um material que auxilie os professores desse nível de ensino no tratamento adequado e correto de conceitos e termos utilizados pelas Ciências Naturais ainda se faz necessário e útil.

Foi nesse sentido que se idealizou o dicionário cuja metodologia de elaboração das definições será aqui descrita. Trata-se de um dicionário destinado a professores do primeiro ciclo do ensino fundamental brasileiro. Antes de descrever a metodologia de elaboração das definições, cumpre destacar que o *corpus* para a seleção de termos e contextos de uso foi composto por coleções didáticas indicadas pelo Ministério de Educação do Brasil (MEC), no âmbito Programa Nacional do Livro Didático (PNLD) de 2010 e 2013, que obtiveram as melhores avaliações no quesito “Conhecimentos e Conteúdos”. Desse modo, as seguintes coleções fazem parte do *corpus* de pesquisa: *A Escola é Nossa* (Editora Scipione); *Caracol - Ciências* (Editora Scipione); *Brasiliiana - Ciências* (Companhia Editora Nacional); *Porta Aberta - Ciências - Edição Renovada* (Editora FTD). Devido à dificuldade de encontrar uma das coleções indicadas pelo PNLD, a *Coleção Pitangua – Ciências*, publicada pela Editora Moderna, foi acrescentada ao *corpus* a *Coleção Buriti – Ciências*, após consulta e indicação da própria editora, como sendo uma coleção equivalente.

1. MAPAS CONCEITUAIS E DEFINIÇÕES: alguns apontamentos teóricos

Haja vista que o objetivo do dicionário é auxiliar os professores, buscou-se, ao selecionar os termos presentes nos livros didáticos, um levantamento efetivo dos conceitos que os docentes devem dominar, assim como o nível de complexidade esperado para cada um deles. Dessa forma, a metodologia seguida neste trabalho segue os pressupostos da Teoria Comunicativa da Terminologia (TCT) uma vez que ele é descritivo e consiste na recopilação de unidades utilizadas por especialistas de uma área em determinada situação de comunicação.

Com isso, entendem-se os termos como:

[...] *unidades lexicais, ativadas singularmente* por suas condições pragmáticas de adequação a um tipo de comunicação [...] a forma é constante; mas o conteúdo se singulariza na forma de seleção de características adequadas a cada tipo de situação e determinadas pela área, pelo tema, pela perspectiva de abordagem do tema, pelo tipo de texto, pelo emissor, pelo destinatário, e pela situação (Cabré, 1999, p. 123).

As unidades terminológicas são, portanto, tidas como poliédricas, uma vez que variam com o tipo de especialista ou especialidade.

Assim como os termos, as definições terminológicas (DT) também são determinadas pela situação comunicativa. Segundo Finatto (2001, p. 161),

[...] a definição, de qualquer tipo, por sua natureza multifacetada, compreende também a interação entre os sujeitos que participam da interlocução que ela instaura e é estabelecida em torno de um com-

portamento linguístico específico que a identifica, modelado por diferentes fatores.

A autora, a partir disso, propõe a separação do texto da DT em dois grupos: o da *significação nuclear* e o da *significação circundante*. O primeiro, que diz respeito ao início da definição, busca responder a seguinte pergunta: “sendo X o termo lematizado, o que é X?”. Já o segundo marca como as informações mais enciclopédicas são incluídas na definição, dependendo das necessidades dos futuros consulentes.

Tendo isso em vista, para a elaboração das primeiras definições do dicionário, foram selecionados termos referentes ao corpo humano. A escolha dos termos foi feita considerando a quantidade de contextos presentes na base de dados, e a relação entre os sistemas do corpo humano, de modo a facilitar a organização e o controle da rede de remissivas. Logo, iniciou-se o trabalho com os termos dos sistemas digestório, urinário, respiratório e genital.

Uma vez separados os termos, eles foram ordenados com base nos níveis de organização do corpo humano - sistema, órgão, tecido e célula - e os seus contextos foram lidos e classificados. De acordo com Barros (2004), há três tipos de contextos que são classificados a partir das informações neles presentes: o *contexto definatório*, que traz características precisas sobre o termo; o *contexto explicativo*, que traz dados sobre o termo, mas sem defini-lo claramente; e o *contexto associativo*, que faz associações auxiliando na indicação da área a qual o termo pertence.

A partir da leitura dos contextos, foi possível perceber que os contextos definitórios e explicativos traziam informações recorrentes sobre os conceitos: o que é, qual a sua função no corpo humano, onde estava localizado e o seu funcionamento. Com isso, pôde-se elencar

quais características são essenciais para os professores e que deveriam constar da DT.

Durante a análise dos contextos, notou-se que a importância da função de uma parte do corpo se dava pela necessidade de explicar o funcionamento de um sistema e seu processo. Isto é, as características de *estômago* são apresentadas a partir do papel do órgão dentro do sistema digestório e da digestão. Sendo assim, foram elaborados mapas conceituais que permitissem a visualização dos processos que acontecem em cada um dos sistemas do corpo humano, além de indicar também o que é cada um dos conceitos.

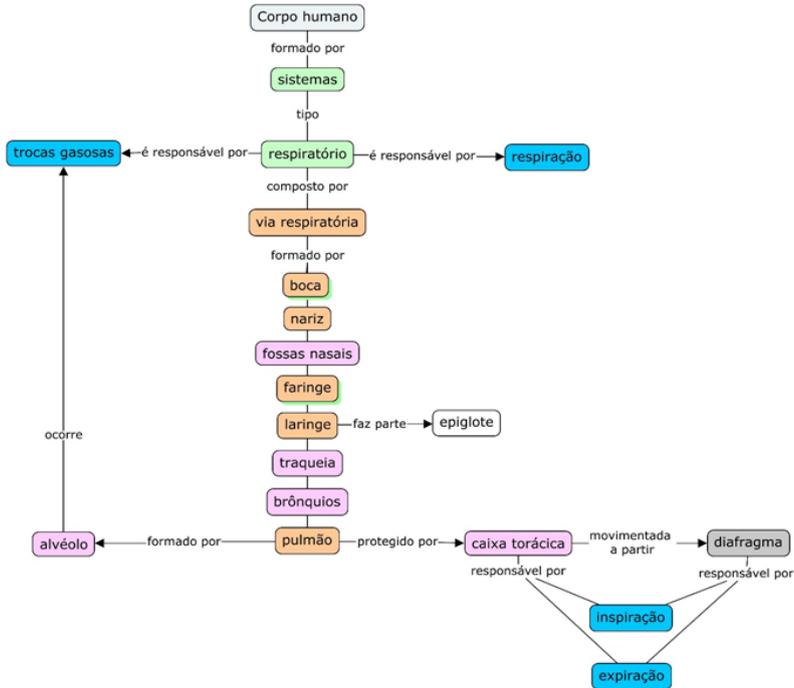
De acordo com Arntz e Picht (1995, p. 104-106), o mapa conceitual deve considerar o intuito de sua elaboração, que área ele sistematiza e quais critérios de organização serão aplicados. Para tal, ele deve seguir alguns princípios: 1) o da univocidade, garantindo a representação, clara e inequívoca, das relações e dos critérios de classificação; 2) o da facilidade de compreensão, assegurando que o nível de complexidade é adequado aos seus usuários; 3) o da transparência, permitindo que os mapas conceituais possam ser decompostos em partes; e 4) o da possibilidade de ampliação, certificando que o mapa conceitual é organizado de maneira flexível e pode receber novas informações.

Seguindo as diretrizes dos autores, foram estruturados quatro mapas conceituais – um para cada sistema do corpo humano – a partir do programa *CmapTools*. A ferramenta permite que se escreva algumas proposições e a partir delas um mapa é apresentado, sendo possível fazer alterações.

Os mapas conceituais elaborados levaram em consideração as informações presentes nos contextos. Dessa forma, utilizaram-se cores para indicar o que é cada um dos conceitos, e as relações entre os

itens foram baseadas na ordem em que o processo (respiração, digestão, reprodução e micção) ocorre.

Figura 1 – Mapa Conceitual do Sistema Respiratório



Fonte: Mota (2019)

O mapa da Figura 1 representa o sistema respiratório e ilustra alguns dos elementos que estão presentes nos demais. Ou seja, as cores, que indicam o que é, são as mesmas em todos os sistemas: verde para sistema; azul para processo; salmão para órgão; rosa para tubos; e cinza para músculos. Além disso, estão com uma borda verde os conceitos que fazem parte de outro(s) sistema(s), como é o caso de *boca* e *faringe* que também atuam no sistema digestório.

A partir das relações evidenciadas no mapa conceitual observa-se o desenvolvimento de todo o processo da respiração, já que ao apresentar a formação da *via respiratória* vê-se as partes do corpo ordenadas respeitando a ordem do processo.

Com a finalização dos mapas conceituais, pôde-se perceber que eles cumprem com o papel de organizar sistematicamente os termos e os conceitos das Ciências Naturais. Além disso, depois de prontos, observou-se como eles também podem ser usados como recurso didático, permitindo a visualização dos conceitos dentro da área de especialidade e suas relações.

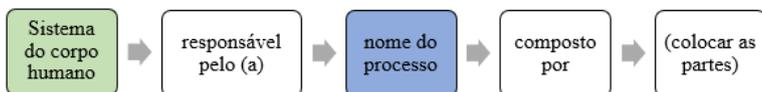
2. ELABORANDO AS DEFINIÇÕES: procedimentos metodológicos

Com os mapas conceituais e com o levantamento das informações mais recorrentes nos contextos, tornou-se possível a elaboração de “diferentes modelos de microestrutura, de modo que sejam constantes para unidades terminológicas de mesma natureza e atendam, ao mesmo tempo, às particularidades dos grupos de termos” (Barros, 2004, p. 200). Os modelos definitórios são padrões de definição que indicam onde deverão constar as informações essenciais para os futuros consulentes.

A partir da análise dos contextos, percebeu-se que as características *o que é e função* são dados relevantes para os professores de Ciências do ensino fundamental I, portanto, são necessários na elaboração das DTs. Já com os mapas conceituais, ficou evidente *o que é*, o que tornou possível indicar nos próprios modelos esta informação, deixando-a indicada com a cor correspondente no mapa conceitual (Figuras 2 e 3).

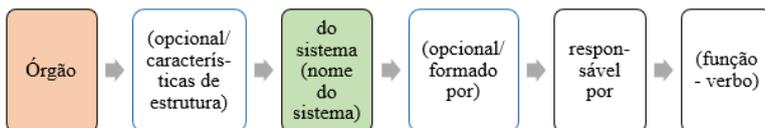
Sendo assim, sete modelos definitórios foram criados auxiliando na escrita das definições de processos, sistemas, órgãos, glândulas, tubos, células sexuais e líquidos:

Figura 2 – Modelo definitório para Sistema



Fonte: Mota (2019)

Figura 3 – Modelo definitório para Órgão



Fonte: Mota (2019)

Os modelos começam sempre com um descritor inicial da mesma classe gramatical do termo-entrada e traz a informação *o que é*, que estava nos mapas conceituais. Há ainda a indicação do sistema ao qual aquele conceito pertence e a sua função sempre indicada por *responsável por* seguida de um verbo. Em alguns casos, é possível acrescentar características de estrutura ou de formação, como é o caso de órgãos externos.

No entanto, a definição não está completa apenas com os modelos definitórios, pois há características nos contextos que não são compreendidas pelos padrões de DT. Para isso, considerou-se as significações de Finatto (2001) em que estariam contempladas na significação nuclear as informações presentes nos modelos, e na significação circundante o que é mais enciclopédico e que varia de acordo com o termo.

Quadro 1 – Significação nuclear do termo *laringe*

SIGNIFICAÇÃO NUCLEAR (sendo X o termo lematizado, o que é X?)

Órgão que recebe o ar da faringe e o envia para a traqueia

Fonte: Mota (2019, p. 93)

Quadro 2 – Significação circundante do termo *laringe*

SIGNIFICAÇÃO CIRCUNDANTE

Na laringe existe a epiglote que funciona como uma tampa direcionando a passagem do ar e dos alimentos

Fonte: As autoras

Somado a isso, considerou-se a lista de práticas apresentadas por Cabré (1993) para a escrita das DTs. Segundo a autora, a DT deve: ser expressa corretamente; utilizar linguagem adequada ao seu público-alvo; constar de uma só frase; utilizar palavras conhecidas, caso contrário, deverão ser entradas no dicionário. Além disso, elas não devem ser circulares, e negações, paráfrases e fórmulas metalinguísticas devem ser evitadas.

Dado o exposto, tem-se a definição de *laringe*.

laringe *s.f.* Órgão do sistema respiratório responsável por permitir a passagem do ar para a traqueia.
Nota: Na parte superior da laringe localiza-se a epiglote.

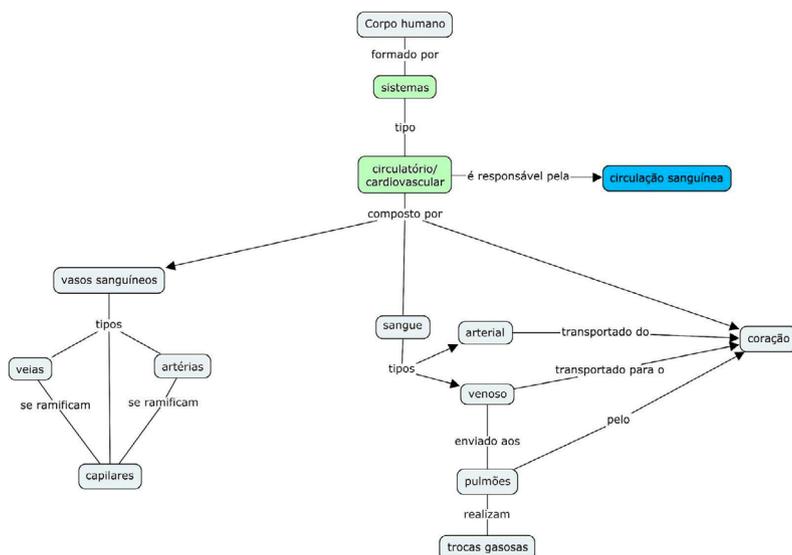
Nela encontra-se a estrutura do modelo definitório com as informações essenciais, seguida de uma **Nota**, em que aparecem dados mais enciclopédicos, mas importantes para os futuros consulentes. Há ainda a indicação da classe de palavra e outros termos das Ciências Naturais estão sublinhados, pois também fazem parte da obra lexicográfica em questão.

3. PADRÕES DEFINITÓRIOS: o sangue é um líquido ou tecido?

Tendo em vista os procedimentos metodológicos descritos, demonstrar-se-á, para exemplificar o seguimento da metodologia, suas vantagens e as dificuldades encontradas na elaboração das definições, os procedimentos adotados na elaboração do termo *sangue*.

Observando o mapa conceitual do sistema circulatório ou cardiovascular, fica evidente a centralidade e a relevância do termo na descrição desse sistema:

Figura 4 – Mapa Conceitual do Sistema Cardiovascular



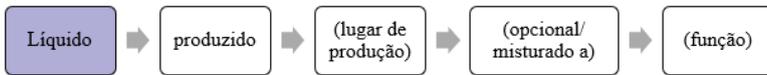
Fonte: as autoras

Conforme a figura, nota-se que o conceito *sangue* foi colocado no centro do mapa, uma vez que todo esse sistema do corpo humano tem como função a circulação desse fluido (*circulação sanguínea*) e sua

transformação, por meio das trocas gasosas. Assim, é essencial que o conceito seja definido com precisão.

A princípio, devido a seu aspecto e recorrendo aos termos genéricos usados anteriormente, foi cogitada a utilização do termo *líquido*. Para isso, foi buscado o modelo definitório para os *líquidos* e termos cujas definições se basearam nesse modelo:

Figura 5 – Modelo Definitório para Líquido



Fonte: Mota (2019)

bile *s.f.* Líquido produzido pelo figado responsável por separar a gordura no intestino delgado.

saliva *s.f.* Líquido produzido pelas glândulas salivares e misturado aos alimentos na boca para dar início à quebra do amido.

Verificou-se, dessa forma, que não seria possível definir *sangue* como um *líquido* seguindo-se tal modelo definitório, uma vez que não seria possível delimitar um “lugar de produção”. Averiguou-se ainda que os líquidos definidos no dicionário referentes ao corpo humano eram, em sua maioria, produzidos por glândulas:

figado *s.m.* Glândula do sistema digestório responsável por produzir a bile.

glândula salivar *s.f.* Glândula do sistema digestório responsável por produzir saliva.

Para além dessas constatações, é necessário apontar que o modelo definatório estabelecido para a definição de *líquidos* demonstrou-se demasiadamente genérico e impedia a descrição da complexidade e da relevância do *sangue* no âmbito do sistema circulatório e do próprio corpo humano, uma vez que esse sistema está presente em todo o corpo.

Assim, embora *líquido* seja um termo associado ao sangue em alguns contextos encontrados nos livros didáticos, em consulta a textos especializados em Biologia, percebeu-se que o termo *sangue* é definido como um *tecido*. Observando também a definição de *tecido* tomada no projeto desenvolvido, verificou-se que responde à complexidade do termo *sangue*.

tecido *s.m.* estrutura formada por um conjunto de células organizadas que atuam de forma integrada para desempenhar determinada função, não estando disposta no corpo em uma única parte isolada.
Nota: No corpo humano há quatro tipos de tecido: epitelial, conjuntivo, nervoso e muscular.

Tendo essas questões em vista, optou-se por uma definição mais precisa e completa, como se poderá verificar a seguir:

sangue *s.m.* tecido do sistema circulatório formado pelo plasma (parte líquida) e por células (hemácias e leucócitos) e fragmentos de célula (plaquetas), responsável por transportar e distribuir os nutrientes, absorvidos pelo sistema digestório, e o oxigênio captado pelo sistema respiratório; também é responsável por retirar o gás carbônico das células e levá-lo aos pulmões, que realizará a troca gasosa.
Nota: Há quatro tipos básicos de sangue humano: A, B, AB e O.

1.: Cf. sangue venoso

2.: Cf. sangue arterial; sangue oxigenado

Com isso, a DT de *sangue* não segue um modelo definitório específico (pois não havia um específico para os tecidos), mas traz as informações essenciais levantadas anteriormente, ou seja, *o que é, de que sistema faz parte* e qual a sua *função*. Assim como as demais DTs elaboradas para o dicionário, a função é introduzida pela expressão *responsável por* seguida de um verbo.

CONSIDERAÇÕES FINAIS

O desenvolvimento deste texto teve como finalidade demonstrar a necessidade de uma obra terminográfica sobre as Ciências Naturais destinado a professores do primeiro ciclo do ensino fundamental e descrever os procedimentos metodológicos seguidos para a definição de termos que comporão um material terminográfico preparado com essa finalidade.

A aplicação desses procedimentos permitiu não apenas a elaboração das definições de forma padronizada, precisa e adequada ao público-alvo, mas também a constatação de que os mapas conceituais podem constituir-se em importantes recursos didáticos, para a visualização das relações que se estabelecem entre os diferentes conceitos da área.

Tal constatação levou a uma reflexão a respeito da forma de apresentação do dicionário, tendo em vista a complexidade da elaboração das definições e a contribuição didática relevante que os mapas conceituais elaborados podem trazer aos futuros consulentes. Assim, decidiu-se que a publicação dos resultados da pesquisa não será reali-

zada em uma única obra terminográfica, mas em obras menores, das quais constarão, além das definições, os mapas conceituais. Acredita-se que tal formato possa contribuir de forma mais significativa aos professores, devido a seu caráter didático.

É válido ressaltar que todo o processo de elaboração das definições foi acompanhado por especialistas das Ciências Naturais, de modo a garantir a qualidade das informações e a sua adequação ao público-alvo. Sendo assim, foram feitas adaptações sempre que necessárias, como foi o caso da DT de *sangue*, o que evidencia, mais uma vez, que os modelos auxiliam na escrita do texto definitório, mas que o objetivo final será sempre a pertinência das características para definir o termo no seu uso. Neste caso, elas terão de ser adequadas aos professores que atuam nas salas de aula de Ciências do ensino fundamental I.

REFERÊNCIAS

ARAÚJO, Mariângela de; SOUZA, Paulo Henrique de. A elaboração de um dicionário terminológico das Ciências Naturais para professores de ensino fundamental I. *In*: CONGRESO INTERNACIONAL DE ENSEÑANZA DE LAS CIENCIAS Y LA MATEMÁTICA - ICIECyM, 1.; ENCUENTRO NACIONAL DE ENSEÑANZA DE LA MATEMÁTICA - II ENEM, 2., 2011. **Actas del** [...]. Tandil: Universidad Nacional del Centro de la Provincia de Buenos Aires, 2011. p. 884-891.

ARNTZ, Reiner; PICHT, Heribert. **Introducción a la Terminología**. Madrid: Fundación Germán Sánchez Ruipérez, 1995.

BARROS, Lidia Almeida. **Curso básico de Terminología**. São Paulo: Editora da Universidade de São Paulo, 2004.

BRASIL. Lei nº 9.394/1996, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional. **Diário Oficial da União**, Brasília, 23 dez. 1996. Disponível em: <https://www2.senado.gov.br/bdsf/handle/id/572694>. [Versão atualizada em 2020]. Acesso em: 05 ago. 2021.

BRASIL. Lei nº 13.005, de 25 de junho de 2014. Aprova o Plano Nacional da Educação (PNE) e dá outras providências. **Diário Oficial [da] República Federativa do Brasil**, Brasília, 26 jun. 2014. Disponível em: https://download.inep.gov.br/publicacoes/institucionais/plano_nacional_de_educacao/plano_nacional_de_educacao_pne_2014_2024_linha_de_base.pdf. Acesso em: 05 ago. 2021.

BRASIL. **Relatório do 1º ciclo de monitoramento das metas do PNE:** biênio 2014-2016. Brasília, DF: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2016.

CABRÉ, Maria Teresa. **La Terminología:** teoría, metodología, aplicaciones. Barcelona: Empúries, 1993.

CABRÉ, Maria Teresa. **La Terminología:** representación y comunicación: Elementos para una teoría de base comunicativa y otros artículos. Barcelona: Institut Universitari de Linguística Aplicada – Universitat Pompeu Fabra, 1999.

FINATTO, Maria José. B. **Definição terminológica:** fundamentos teórico-metodológicos para sua descrição e explicação. 2001, 395f. Tese (Doutorado em Estudos da Linguagem, Teorias do Texto e do Discurso) – Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2001.

MOTA, Ariane Vicente. **A elaboração de modelos definitórios para um Dicionário Terminológico das Ciências Naturais.** 2019. 108 f. Dissertação (Mestrado) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2019.

FRASEMAS SEMÂNTICO-LEXICAIS COMPOSICIONAIS NO ÂMBITO DA TEORIA SENTIDO-TEXTO: ESTUDO DE ALGUNS CLICHÊS DO LÉXICO DO PORTUGUÊS BRASILEIRO

Sabrina Pereira de Abreu
Sara Augusto Carra

INTRODUÇÃO

Este trabalho constitui-se em um exercício exploratório acerca de um tipo especial de expressões fraseológicas (ou frasemas): os clichês, comumente entendidos como ‘*frase que se banaliza por ser muito repetida; lugar-comum, chavão*’ (cf. verbete CLICHÊ, Houaiss [2011]). Frases ou ditos populares como *Minha vida é um livro aberto; Vivendo e aprendendo; Deus escreve certo por linhas tortas; Em casa de ferreiro, espeto de pau; Filho de peixe peixinho é; Não cuspa no prato em que comeu; Quem ama o feio bonito lhe parece; Quem vê cara não vê coração; Minha vida é um livro aberto*; entre tantas outras, são alguns dos clichês populares que “andam de boca em boca” entre os falantes do português brasileiro. Há ainda segmentos (alguns com valor frasal) que se apresentam de forma banalizada, tendo em vista a alta frequência de seu uso: [guardar segredo] *a sete chaves*; [ter uma] *amarga decepção*; [ter uma] *calorosa recepção*; [ficar] *cego de amor*; [fazer uma] *crítica construtiva*; [desencadear

um] *efeito dominó*; [dar] *de mão beijada*; [ser] óbvio e ululante; [fechar com] *chave de ouro*; [fazer] *das tripas coração* etc.

Estas frases e segmentos são de difícil classificação no âmbito dos estudos lexicais, tanto em função da extensão diversificada que apresentam (como vimos, podem ser frases inteiras, um sintagma, uma única palavra etc.) quanto em função do tipo de sentido que veiculam, pois não são opacos; entretanto, seu sentido só se atualiza em contextos pragmáticos específicos.

“Comendo pelas bordas”, neste trabalho, pretendemos, ainda de forma preliminar, examinar um conjunto de unidades fraseológicas do léxico do português brasileiro inventariadas como “lugares-comuns” ou “frases feitas” colhidas de um dicionário organizado por Werneck (2014) durante mais de três décadas. O autor não indica nenhum referencial teórico que o tenha guiado na compilação dessas expressões. Apenas informa que, durante “quase quatro décadas”, como um colecionador, foi recolhendo essas expressões, que acabaram sendo organizadas em ordem alfabética em um dicionário (Werneck, 2014, p. 8). Considerando a diversidade de unidades fraseológicas listadas na obra, propomo-nos a examiná-las do ponto de vista da Teoria Sentido-Texto (daqui em diante, TST), que inclui essa sorte de unidade fraseológica entre os tipos de frasemas: os clichês.

Para “trilhar essa seara”, organizamos o trabalho da seguinte maneira: na seção 2, apresentamos a discussão teórica acerca do que se entende por frasema no escopo da TST; na seção 3, tratamos especificamente dos frasemas semântico-lexicais composicionais (os clichês) na perspectiva da teoria que guiará a análise dos dados; na seção 4, apresentamos os procedimentos metodológicos adotados na pesquisa; na seção 5, mostramos a análise dos dados e os resultados obtidos; e, por fim, seguem as considerações finais.

1. TIPOS DE FRASEMAS SEGUNDO A TST: locução, colocação e clichês

Para Mel'čuk (2013), frasemas são expressões fraseológicas (ou multilexêmicas) não livres, ou seja, são construídos com, no mínimo, dois lexemas e apresentam restrições de seleção no eixo paradigmático e de combinatória no eixo sintagmático. Isto significa que o Locutor escolhe o sentido que quer expressar e, entre as possibilidades de paráfrases possíveis para expressar esse sentido, elege uma expressão fraseológica. Por exemplo, para expressar o sentido 'chover muito <intensamente>', o Locutor pode escolher a sequência *a cântaros*, que é uma unidade lexical multilexêmica que tem valor intensificador (chover 'a cântaros' significa chover 'muito, intensamente'). Polguère (2015, p. 5) ensina-nos que, no âmbito da TST, há três tipos de expressões fraseológicas (ou frasemas), são elas: as locuções, as colocações e os clichês linguísticos.

As locuções (p. ex., [coração] DE ANJO (= angélico); BATER AS BOTAS (= morrer)) distinguem-se das colocações e dos clichês por constituírem unidades lexicais sintagmáticas e, portanto, terem, no sistema linguístico, a mesma autonomia que os lexemas. Nesta perspectiva, locuções são expressões fraseológicas não composicionais, ou seja, resultam de um todo semântico: os elementos que constituem esse tipo de frasema não atualizam seus sentidos para formar esse todo semântico. Polguère (2015, p. 6) diz ainda que a distinção entre locuções e lexemas "[...] está na sua complexidade de forma, pois ela possui uma estrutura interna de natureza sintática", ressaltando-se que os sentidos veiculados por seus elementos constitutivos são opacos.

O segundo tipo de frasema são as colocações. Conforme Mel'čuk (2013, p. 138), as colocações são frasemas lexicais semi-res-

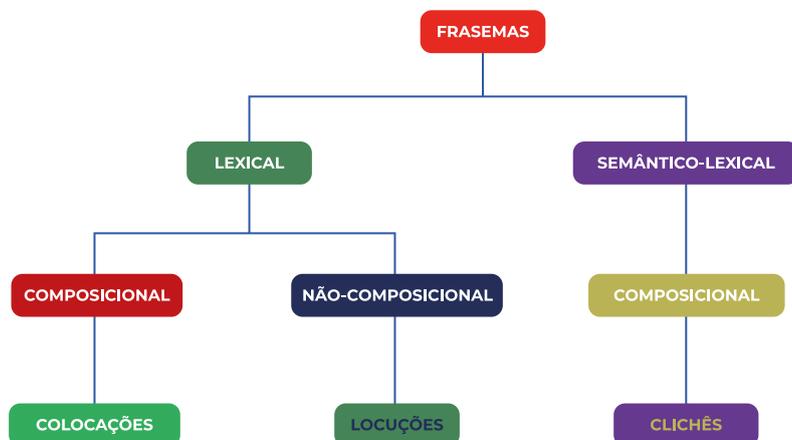
tritros, pois um de seus componentes, a base da colocação, é selecionado livremente pelo Locutor, enquanto o outro elemento, chamado “colocado”, é selecionado em função do sentido veiculado pelo primeiro elemento. Como exemplos de colocações no português, temos: CORRER rapidamente; CHOVER a cântaros. Abreu (2018, p. 5), assim, sintetiza as propriedades desse tipo de frasema:

*A colocação, por sua vez, é uma combinação frequente de duas ou mais palavras em um mesmo contexto, onde a seleção de um item lexical desencadeia a ocorrência de outro. Uma colocação pode ser dividida em duas partes: a base e o colocado, que correspondem ao número de lexias que uma colocação contém. Por exemplo, em *boa colheita*, *boa* é o colocado, e *colheita*, a base da colocação; em *dar um golpe*, *dar* é a base, e *golpe*, o colocado; em *amar loucamente*, *amar* é a base, e *loucamente*, o colocado). Os frasemas colocacionais podem ser de diferentes tipos: (a) verbo de suporte + complemento (*levar [em consideração] = considerar*); (b), substantivo + adjetivo (*leite magro*); (c) verbo pleno + segmento com valor adverbial (*trabalhar [como uma mula]*), etc. (Abreu, 2018, p. 5).*

Por fim, o último tipo de frasema é conhecido como “clichês linguísticos” (p. ex., *proibido estacionar; caixinha de surpresas; agradecer a gregos e troianos*), que são frasemas semântico-lexicais composicionais. Isso quer dizer que, sendo composicional, as unidades lexicais que compõem o clichê contribuem para o sentido total da expressão; e, sendo semântico-lexical, as unidades lexicais que compõem o clichê, bem como o seu sentido, são selecionadas de maneira restrita pelo conteúdo conceitual a ser expresso, isto é, a situação de comunicação interfere na seleção da expressão.

Nosso exercício analítico, concentra-se nos frasemas semântico-lexicais composicionais, ou seja, unidades formadas por, pelo menos, dois lexemas ligados sintaticamente, cujos elementos constitutivos e sentido são selecionados pelo Locutor de forma restrita, e o sentido do todo é entendido pela soma dos sentidos dos elementos constitutivos. A seguir, com base em Mel'čuk (2012, p. 37), resumimos, em uma figura, os tipos de frasemas e suas características.

Figura 1 – Tipos de frasemas e suas características



Fonte: Elaborado pelas autoras a partir de Mel'čuk (2012, p. 37)

2. FRASEMAS SEMÂNTICO-LEXICAIS COMPOSICIONAIS: clichês linguísticos

Como vimos na seção anterior, os frasemas podem ser não-composicionais, como as locuções, ou composicionais, como as colocações e os clichês (Mel'čuk, 2012, p. 1). Os clichês, objeto de análise neste trabalho, são frasemas semântico-lexicais composicionais, ou seja, unidades formadas por, pelo menos, dois lexemas ligados sintaticamente.

ticamente, cujos elementos constitutivos e sentido são selecionados pelo Locutor de forma restrita, e o sentido do todo é entendido pela soma dos sentidos dos elementos constitutivos.

Em seu texto intitulado *Phraseology in the language, in the dictionary, and in the computer*, publicado em 2012, Mel'čuk afirma que os clichês linguísticos são “[...] os mais numerosos entre os frasemas, são também os mais difíceis de localizar e coletar” (Mel'čuk, 2012, p. 40). O autor define o clichê linguístico como “[...] uma expressão composicional [...] usada para expressar um sentido complexo ‘ σ ’ que uma língua L prescreve para ser usada na descrição de uma dada situação P [...]” (Mel'čuk, 2012, p. 40).

O autor afirma que esse tipo de frasema é caracterizado por uma âncora lexical (ou âncoras), que é o lexema cujo significado identifica o uso do clichê. Por exemplo, na pergunta-clichê *Qual é o seu nome?*, o lexema NOME é identificado como a âncora do clichê, a qual não precisa . Em um dicionário, os clichês são descritos nos verbetes de suas âncoras lexicais.

Os clichês são divididos em dois grupos:

Grupo 1 - pragmatemas, que sofrem restrições pragmáticas – fórmulas exigidas em uma situação particular de uso (*aguarde na linha*); e

Grupo 2 - normais, que não sofrem restrição pragmática– como os ditos populares (*mentira tem perna curta*) e os nomes próprios complexos (*Novo/Velho Testamento*).

O Grupo 1, que se define pela restrição pragmática, necessita de uma situação particular de uso para que seu significado possa ser compreendido. Os clichês deste grupo podem se apresentar por meio de quatro subtipos, como se observa no quadro abaixo.

Quadro 1 – Pragmatemas

Clichês = pragmatemas (pragmaticamente restritos)			
Subtipo 1: Colocação	Contexto situacional	Subtipo 2: Frase	Contexto situacional
<i>Tinta molhada</i> <i>Tinta fresca</i>	[aviso em uma parede recém pintada]	<i>Segure a linha!</i> <i>Ênfase minha/nossa).</i> <i>Devolver ao remetente.</i> <i>Quem está aí?</i>	[em uma conversa telefônica] [após uma citação em um texto escrito] [em uma carta enviada pelo correio] [após uma batida na porta]
Subtipo 3: Locução	Contexto situacional	Subtipo 4: Lexema-	Contexto situacional
[olhos] <i>da cor do mar</i>	[para designar a cor de algo]	<i>Merda!</i>	[para um artista que vai ao palco]

Fonte: Elaborado pelas autoras (a partir de Mel'čuk, 2012)

O Grupo 2 é também conhecido como “clichê pragmaticamente não restrito”. Nele, o sentido linguístico é o sentido literal do clichê, ou seja, corresponde ao seu conteúdo informacional, expressa o estado de coisas que o clichê por si mesmo denota. Para usá-lo corretamente, o Locutor precisa saber que o clichê é usado para expressar determinado conceito. Tomemos como exemplo dos clichês do Grupo 2 o provérbio *Água mole em pedra dura tanto bate até que fura*. Para usá-lo corretamente, o Locutor precisa saber que esse provérbio expressa o seguinte significado: ‘Se você não parar de insistir, de tentar, acabará alcançando o que deseja’.

Os clichês do Grupo 2 também se subdividem. Essa subdivisão se dá em dois subtipos: provérbios e nomes próprios. Nestes últimos, o sentido complexo é também literal, mas aqui temos um referente particular, como ocorre, por exemplo, quando usamos o clichê *Planeta azul* para nos referirmos ao planeta Terra.

O quadro a seguir ilustra os dois tipos de clichês pragmaticamente restritos.

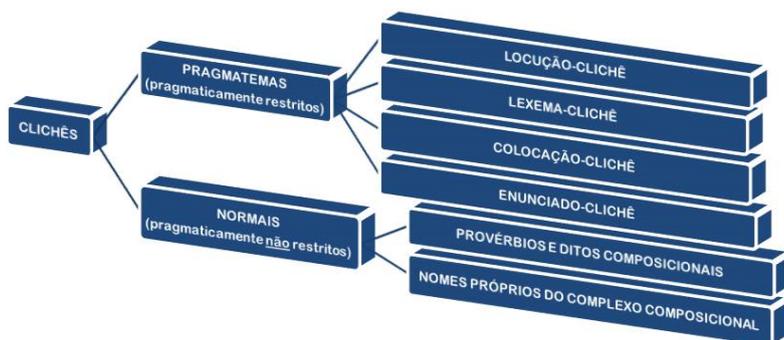
Quadro 2 – Clichês normais

Clichês normais (pragmaticamente não restritos)	
Subtipo 1: Provérbios e ditos composicionais	Subtipo 2: Nomes próprios do complexo composicionais
<i>O dinheiro não é tudo.</i>	<i>Velho/Novo Testamento</i>
<i>Deus ajuda quem madruga.</i>	<i>Sonho de uma Noite de Verão</i> [peça de Shakespeare]
<i>Não adianta chorar pelo leite derramado.</i>	<i>Cidade Luz</i> [Paris]
<i>A pressa é inimiga da perfeição.</i>	<i>Cidade Eterna</i> [Roma]
<i>As aparências enganam.</i>	<i>planeta vermelho</i> [Marte]
<i>Cada macaco no seu galho.</i>	<i>paneta azul</i> [Terra]

Fonte: Elaborado pelas autoras (a partir de Mel'čuk, 2012)

Em síntese, os clichês linguísticos podem ser condicionados pragmaticamente ou não. Quando condicionados pela situação de uso, são conhecidos no âmbito da TST como pragmatemas e podem se apresentar com diferentes extensões. Já os clichês ditos “normais” são lexemas, locuções, colocações ou frases que são usadas com muita frequência pelos Locutores e que se tornaram desgastadas, comuns. A figura abaixo resume a tipologia dos clichês que apresentamos nesta seção.

Figura 2 – Tipos de clichês



Fonte: Elaborada pelas autoras

Com base nessa classificação, pretendemos verificar quais são os tipos de clichês inventariados na obra supracitada. O objetivo desse exercício explanatório é contribuir com as discussões acerca desse

tipo de frasema no âmbito da TST. Já há trabalhos que descrevem os clichês de acordo com essa teoria, como os de Papadopoulou (2013, 2014) para o grego, mas, até onde pudemos averiguar, não há descrições de clichês do português brasileiro com base nesse quadro teórico.

Certamente, já esperamos de antemão que alguns dos tipos de clichês pragmaticamente restritos, como os lexemas-clichês, e dos pragmaticamente não restritos, como os nomes próprios, não sejam encontrados em nosso *corpus* de análise. Antes de apresentar a análise propriamente dita, na próxima seção, exporemos os procedimentos metodológicos adotados na pesquisa.

3. PROCEDIMENTOS METODOLÓGICOS

As unidades foram coletadas do livro *O pai dos burros – dicionário de lugares comuns e frases feitas*, cujo autor, Humberto Werneck, é importante ressaltar, não é linguista, mas um jornalista nascido em Minas Gerais (MG).

Figura 3 – Capa de Werneck (2014)



Fonte: Werneck (2014) - capa

O livro arrola diversas unidades que o autor entende como *clichês, lugares comuns e/ou frases feitas* e foi escrito – aparentemente, pelo menos – para escritores, pois ele afirma:

[...] se escrever vale a pena, deve ser para enunciar algo que se pretende novo – e me parece um contrassenso, sobretudo no jornalismo, tentar passar o novo com linguagem velha. [...]

[...] O que se quer com esse livro é apenas recomendar desconfiança diante de tudo aquilo que, no ato de escrever, saia pelos dedos com demasiada facilidade. Porque nada de verdadeiramente bom costuma vir nesse automatismo. [...] (Werneck, 2014, p. 10-11)

O autor entende *clichês* como *fórmulas prontas* usadas por quem tem medo de pensar um modo novo de se querer expressar uma ideia, ou seja, o que o senso comum entende por clichês e frases feitas, assim, evidentemente, não há qualquer teoria lexicográfica ou mesmo linguística por trás desse livro. Sua coleta de dados se deu através de clichês que ele (ou seus amigos e colegas de redação de jornais e revistas) encontrava aleatoriamente, anotava e guardava em um envelope durante quase quatro décadas, segundo o próprio autor.

Se entendermos o dicionário como um repositório do léxico e se entendermos que existem diversos tipos de dicionários, que apresentam as mais diversas estruturas e os mais diversos recortes lexicais, talvez, poderíamos considerar o livro como dicionário. Entretanto, como afirma Biderman (1998, p. 129), “[...] o dicionário constitui uma organização sistemática do léxico, uma espécie de tentativa de descrição do léxico de uma língua” e, se observarmos a figura abaixo, veremos que o ‘dicionário’ não nos fornece qualquer tipo de descrição, se assemelhando muito mais a uma listagem de clichês (no enten-

dimento do autor sobre o que seja *clichê*), cujas entradas não se pode entender se se trata da primeira palavra da unidade ou se é o que ele entende como núcleo, como podemos ver na entrada ‘**chute**’ (*chute certo* e o *chute não oferece perigo (fut.)*) e na entrada ‘**cintura**’ (*ter jogo de cintura*); o autor não nos informa como ele organizou essas entradas. Além disso, há apenas uma marca de uso no livro, que é o uso *futebolístico*, que ele marca como (**fut.**), não há qualquer outra marca, nem pragmática (como *uso informal*), nem linguística (não há marcas de locução, nem de colocação, nem de clichê, nem, no caso das locuções, a classe gramatical – como locução verbal ou locução nominal) e também não apresenta qualquer tipo de definição:

Figura 4 – Página de Werneck (2014)

Chute	Chute certo O chute não oferece perigo (fut.)
Chuteira	A pátria de chuteiras, como disse Nelson Rodrigues Pendurar as chuteiras
Chuva	A chuva deu/não deu trégua A chuva transformou as ruas em verdadeiros rios Castigado pela chuva Chuva a tamborilar no telhado Chuva copiosa Chuva criadeira Chuva fina e incessante Chuva miúda Chuva que castiga Chuva torrencial Grossas bategas de chuva Quem está na chuva é para se molhar Sujeito a chuvas e trovoadas
Cidadania	Promoção da cidadania
Cidadão	Cidadão comum Cidadão de bem Cidadão exemplar Cidadão pacato
Ciência	Nem a ciência pode explicar
Cifras	Cifras astronômicas
Cinema	Coisa de cinema
Cinto	Apertar o cinto
Cintura	Ter jogo de cintura

Fonte: Werneck (2014)

O próprio autor, na introdução, assume que não se trata de um dicionário, mas, sim, de um mero repertório. O autor, conjuntamente com seus colegas e amigos, catalogaram durante mais de três décadas, unidades de sentido que **percebiam** como muito utilizadas pelas pessoas.

O dicionário registra **5.333 unidades** entendidas como frases feitas ou lugares comuns pelo autor. A coleta dessas unidades foi feita manualmente. Foram excluídas **141 unidades** que receberam a marca (*fut.*), que assinala registros usados no âmbito do *futebol* (como podemos ver na Figura 4: *O chute não oferece perigo (fut.)*) e **661 unidades** que não são frasemas, mas frases livres (como podemos ver na imagem acima: *A chuva transformou as ruas em verdadeiros rios*). Assim, foram analisadas **4531 unidades**.

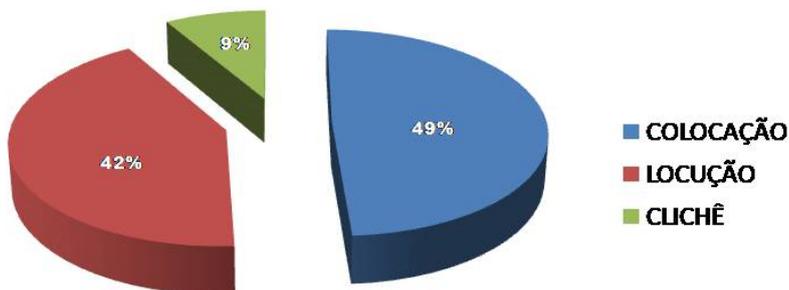
A partir disso, a análise ocorreu em quatro etapas: (1) verificamos que **tipos de frasemas** foram encontrados no dicionário; (2) verificamos que **tipos de clichês** foram encontrados no dicionário; (3) verificamos que tipos de **pragmatemas** foram encontrados no dicionário; e, por fim, (4) verificamos que tipos de **clichês normais** foram encontrados no dicionário.

4. ANÁLISE DOS DADOS E RESULTADOS PRELIMINARES

4.1. Tipos de frasemas encontrados em Werneck (2014)

Foram encontradas **1920 locuções** (o que corresponde a **42%** do *corpus* analisado), **2223 colocações** (o que corresponde a **49%** do *corpus* analisado) e **388 clichês** (assim, nosso *corpus* de pesquisa corresponde a **9%** dos frasemas identificados).

Gráfico 1 – Tipos de frases encontradas em Werneck (2014)



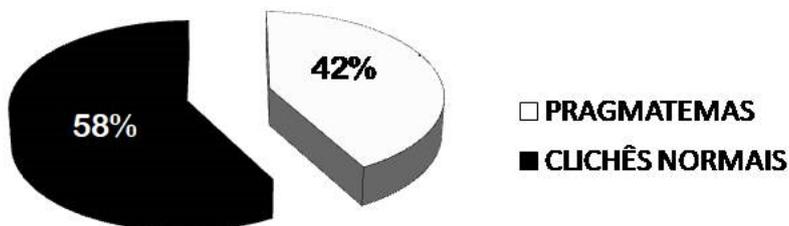
Fonte: Elaborado pelas autoras

Alguns exemplos de **locuções** encontradas são: *descascar o abacaxi; abraço de tamanduá; amigos do alheio; ter asas no pé; dar bandeira; sem pé nem cabeça; fazer a caveira de alguém*. Alguns exemplos de **colocações** encontradas são: *agressão brutal; profundamente abalado; estudar/trabalhar com afinco; trágico acidente; acidente fatal; completo abandono; afirmar categoricamente*. Alguns exemplos de **clichês** encontrados são: *tudo de bom!; quem cala consente; bondade sua; em briga de marido e mulher, ninguém mete a colher; vira essa boca pra lá; nem por cima do meu cadáver; todos os caminhos levam a Roma*.

4.2. Tipos de clichês encontrados em Werneck (2014)

Foram encontrados **163 pragmatemas** (o que corresponde a 42% dos clichês analisados) e **225 clichês normais** (o que corresponde a 58% dos clichês analisados).

Gráfico 2 –Tipos de clichês encontrados em Werneck (2014)



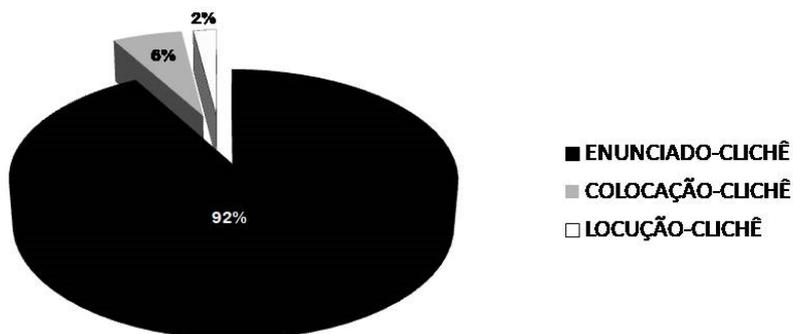
Fonte: Elaborado pelas autoras

Alguns exemplos de **pragmatemas** encontrados são: *vá pentear macaco!*; *é só uma lembrancinha*; *qual é a sua graça?*; *é importante frisar*; *como vai essa força?*; *o dever me chama*; *voltamos depois do intervalo*. Alguns exemplos de **clichês normais** encontrados são: *contra fatos, não há argumentos*; *o Velho Mundo*; *para morrer, basta estar vivo*; *a necessidade é a mãe da invenção*; *o inferno está cheio de boas intenções*; *a primeira impressão é a que fica*; *mais vale um pássaro na mão do que dois voando*.

4.3. Tipos de pragmatemas encontrados em Werneck (2014)

Entre os **pragmatemas** encontrados, **92%** correspondem a **enunciados-clichê**, **6%** correspondem a **colocações-clichê** e **2%** correspondem a **locações-clichê**. Não foi encontrado **lexema-clichê**.

Gráfico 3 – Tipos de pragmatemas encontrados em Werneck (2014)



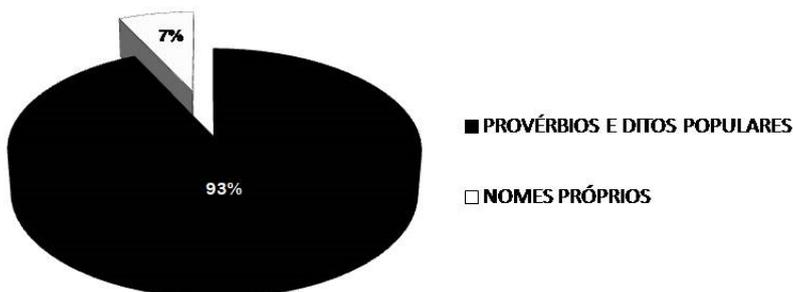
Fonte: Elaborado pelas autoras

Um exemplo de **locução-clichê** encontrado é *par de jarros*. Alguns exemplos de **colocação-clichê** encontrados são: *honra ao mérito; orgulhosamente apresenta; antecipadamente agradece; meus sentimentos; triste sina; sinceros agradecimentos; paz aos mortos*. Alguns exemplos de **enunciado-clichê** encontrados são: *Inês é morta; ou vai ou racha; é pegar ou largar; quem é vivo sempre aparece; faço minhas suas palavras; tenha a santa paciência; é só uma lembrancinha*.

4.4. Tipos de clichês normais encontrados em Werneck (2014)

Entre os **clichês normais** encontrados, 93% correspondem a **provérbios e ditos populares** e 7% correspondem a **nomes próprios**.

Gráfico 4 – Tipos de clichês normais encontrados em Werneck (2014)



Fonte: Elaborado pelas autoras

Alguns exemplos de **provérbios e ditos populares** encontrados são: *para bom entendedor, meia palavra basta; Deus é brasileiro; enquanto há vida, há esperança; nem tudo que reluz é ouro; a paciência é uma virtude; a mentira tem pernas curtas; devagar se vai ao longe*. Alguns exemplos de **nomes próprios** encontrados são: *Guerra Fria; o Bom Velhinho (Papai Noel); Idade das Trevas; Pedra Filosofal; Notório Saber; o Velho Mundo; a pomba da paz*.

CONSIDERAÇÕES FINAIS

Este é um estudo de **caráter preliminar e exploratório**, portanto, para que se possa descrever realmente os clichês do português brasileiro (PB), é necessário fazer um levantamento maior de dados em *corpora* da língua em uso – que pretendemos fazer futuramente. Entretanto, os dados encontrados nos permitem tecer algumas considerações:

- O autor, enquanto Locutor, não distingue as unidades entre tipos de frasesmas, uma vez que insere no mesmo grupo colocações, locuções e clichês, chamando-os de “frases feitas”

ou “lugares comuns”. Isso revela que a direcionalidade pragmática, como um todo, não foi distinguida pelo autor – exceto, talvez, pelas ocorrências distinguidas pelo autor com a marca (*fut.*), uma marca terminológica.

- Entre os **pragmatemas**, não há ocorrência de lexemas. Parece-nos que realmente esse tipo de clichê é difícil de ocorrer no PB. Na verdade, pensamos que há um **pseudo** clichê lexemas, pois em *palavras...*, exemplo recolhido do dicionário analisado, temos claramente apagamento: a unidade seria *são só palavras*, mas o autor registrou apenas *palavras...*
- No que se refere aos **clichês normais**, dados apontam que a imensa maioria é de **provérbios**, não de nomes próprios. Neste caso, parece-nos que os nomes próprios são mais difíceis de serem percebidos pelo Locutor como um clichê – ao contrário dos provérbios e ditos populares.

REFERÊNCIAS

ABREU, Sabrina Pereira de. Considerações sobre a natureza colocacional e locucional de sintagmas terminológicos. **Revista Digital Internacional de Lexicologia, Lexicografía y Terminología**, n. 1, 2018. Disponível em: <https://revistas.unc.edu.ar/index.php/ReDILLeT/issue/view/1746>. Acesso em: 24 ago. 2023.

BIDERMAN, Maria Tereza Camargo. Os dicionários na contemporaneidade: arquitetura, métodos e técnicas. In: OLIVEIRA, Ana Maria Pinto de; ISQUERDO, Aparecida Negri (orgs.). **As ciências do léxico**. Lexicologia, Lexicografia e Terminologia. v. 1. Campo Grande: Editora UFMS, 1998. p. 129-142.

HOUAISS, Antônio. **Dicionário Eletrônico Houaiss da Língua Portuguesa**. São Paulo/Rio de Janeiro: Objetiva, 2009.

MEL'ČUK, Ian Alexander. Phraseology in the language, in the dictionary, and in the computer. **Yearbook of Phraseology**, Berlim, v. 3, p. 31-56, nov. 2012. Disponível em: <https://www.degruyter.com/document/doi/10.1515/phras-2012-0003/html>. Acesso em: 07 out. 2021.

MEL'ČUK, Ian Alexander. Tout ce que nous voulions savoir sur les phrasèmes, mais.... **Cahier de Lexicologie**, Paris, n. 102, p. 129-149, 2013.

PAPADOPOULOU, Lena. “My deepest condolences”: Lexical functions of Greek pragmatemes [inafuneral]. *In*: INTERNATIONAL CONFERENCE ON GREEK LINGUISTICS, 11., 26-29 set. 2013, Rhodes, Grécia. **Proceedings** [...]. Rhodes: University of the Aegean 2013.

PAPADOPOULOU, Lena. One Lexicological Theory, two Lexicographical Models and the Pragmatemes. *In*: EURALEX INTERNATIONAL CONGRESS: THE USER IN FOCUS, 16., 2014. **Proceedings** [...]. Bolzano: Eurac, 2014. p. 1039-1048.

POLGUÈRE, Alain. Non-compositionnalité: ce sont toujours les locutions faibles qui trinquent. **Verbum**, Nancy, n. 2, p. 257-280, 2015. Disponível em: <https://hal.archives-ouvertes.fr/hal-01378528/document>. Acesso em: 08 out. 2021.

WERNECK, Humberto. **O pai dos burros** – dicionário de lugares-comuns e frases feitas. Porto Alegre: Arquipélago Editorial, 2014.

OS AUTORES

Aderlande Pereira Ferraz

Doutor em Linguística e Língua Portuguesa pela Universidade Estadual Paulista Júlio de Mesquita Filho (1998). Pós-doutorado pela Universidade de São Paulo (2022, 2015, 2005). Atualmente é professor associado e docente permanente do Programa de Pós-Graduação em Estudos Linguísticos da Universidade Federal de Minas Gerais (UFMG). Área de atuação: Teoria e Análise Linguística, atuando no ensino de graduação e de pós-graduação, principalmente nas seguintes subáreas: Lexicologia e Lexicografia, com os seguintes temas: descrição e ensino do léxico, neologismos, fraseologismos, lexicografia pedagógica. ORCID: <https://orcid.org/0000-0003-1874-7113>.

Álvaro David Hwang

Doutor e mestre em Letras pela Universidade de São Paulo. Pós-Doutor pela UNESP/São José do Rio Preto (novembro de 2022). Professor de Língua Francesa na Universidade Estadual de Maringá, Paraná. Pesquisador na área de lexicografia geral e lexicografia pedagógica, com pesquisas voltadas à lexicografia de língua francesa. ORCID: <https://orcid.org/0000-0003-4320-8306>.

Ana Paula Tribesse Patrício Dargel

Doutora em Linguística e Língua Portuguesa pela Universidade Estadual Paulista Júlio de Mesquita Filho (2011), Estágio Pós-Doutoral em Linguística pela Universidade Federal de Mato Grosso do Sul (UFMS/CPTL). É professora adjunta na Universidade Estadual de Mato Grosso do Sul, atuando nos Programas de Mestrado Profissional em Letras em Rede - PROFLETRAS e Mestrado Acadêmico em Letras. Tem experiência na área de Linguística e desenvolve pesquisas nos seguintes temas: Lexicologia, Lexicografia, Lexicografia Pedagógica, Sociolinguística Educacional e Onomástica (Toponímia e Antroponímia). ORCID: <https://orcid.org/0000-0002-2197-5259>.

Anna Maria Becker Maciel

Doutora em Letras pela Universidade Federal do Rio Grande do Sul (UFRGS, 2001). Docente aposentada do Instituto de Letras (UFRGS) e do Programa de Pós-Graduação em Letras (UFRGS). Pesquisadora na área de Linguística, atuando em Terminologia, Terminografia e Linguística de Corpus. Co-fundadora do Projeto TERMISUL. CV: <http://lattes.cnpq.br/4727551751956552>.

Aparecida Negri Isquerdo

Doutora em Linguística e Língua Portuguesa pela UNESP/Araraquara (1996). Professora emérita da Universidade Federal de Mato Grosso do Sul (2023) e Pesquisadora Sênior nos Programas de Pós-Graduação em Letras e em Estudos da Linguagem da UFMS. Pesquisadora na área de Linguística, com pesquisas nas áreas de Lexicologia, Lexicografia e Onomástica. Uma das organizadoras da coleção *As Ciências do Léxico* (vol. I, II, III, IV, V, VI, VII, VIII e IX). ORCID: <https://orcid.org/0000-0003-1129-5775>.

Ariane Vicente Mota

Mestra em Letras pela Universidade de São Paulo (2019). Doutoranda em Letras pela mesma instituição e pesquisadora da área de Terminologia, elaborando e analisando definições terminológicas para termos das Ciências Naturais. ORCID: <https://orcid.org/0000-0002-6873-3740>.

Celina Márcia de Souza Abbade

Doutora (2003) e Mestre (1998) em Letras pela Universidade Federal da Bahia. Licenciada em Letras Vernáculas com Francês (1992) e Bacharel em Psicologia (2023) pela UCSAL. Pós-doutora em Estudos de Linguagens (2016) pela UEFS. Professora Plena da UNEB e docente permanente do PPGEL/UNEB e do Colegiado de Letras Vernáculas (DCH-I). Líder do Núcleo de Estudos Lexicais (NEL), coordenando os Projetos Grupo de Estudos Lexicais e Terminológicos da Bahia (GELTBA) e Atlas Toponímico da Bahia (ATOBABH). Uma das organizadoras da coleção *As Ciências do Léxico* (vol. VIII). ORCID: <https://orcid.org/0000-0001-6485-277X>.

Claudia Zavaglia

Doutora em Linguística e Língua Portuguesa pela Universidade Estadual Paulista Júlio de Mesquita Filho (2002). É Livre-Docente em Lexicografia e Lexicologia desde 2009. Atualmente é professora associada da Universidade Estadual Paulista Júlio de Mesquita Filho - UNESP - Campus de São José do Rio Preto - SP - IBILCE. Desde 2003, atua na Pós-graduação em Estudos Linguísticos da UNESP/IBILCE, na linha Lexicologia e Lexicografia, direcionando suas pesquisas, substancialmente, para os estudos lexicográficos, lexicológicos e tradutológicos. ORCID: <https://orcid.org/0000-0003-0250-7019>.

Cleci Bevilacqua

Doutora em Linguística Aplicada pelo Instituto Universitário de Linguística Aplicada, Universidade Pompeu Fabra, Barcelona (2004). Professora convidada do Departamento de Línguas Moderna e do programa de Pós-Graduação em Letras da UFRGS, na linha de pesquisa Estudos do Léxico e da Tradução. Desenvolve pesquisa nas áreas de Terminologia, Fraseologia Especializada e Tradução e suas interfaces com a Linguística de Corpus. É pesquisadora do Grupo de Pesquisa Termisul, líder do Grupo de Pesquisa Termisul no CNPq e pesquisadora do CNPq (PQ2). ORCID: <https://orcid.org/0000-0002-1002-9080>.

Clotilde de Almeida Azevedo Murakawa

Doutora em Letras, Linguística e Língua Portuguesa pela Universidade Estadual Paulista Júlio de Mesquita Filho, Araraquara (1991). Atualmente é professora permanente no Programa de Pós-graduação em Linguística e Língua Portuguesa da Universidade Estadual Paulista Júlio de Mesquita Filho. Tem experiência na área de Letras, com ênfase em Língua Portuguesa, atuando principalmente nos seguintes temas: lexicografia, lexicologia, linguística histórica e semântica. ORCID: <https://orcid.org/0000-0002-3815-3534>.

Elizabete Aparecida Marques

Doutora em Linguística Aplicada pela Universidad de Alcalá (Espanha, 2007). Estágio Pós-Doutoral na área da Fraseologia pela Université Sorbonne Paris Nord (França, 2013). Professora Titular na Universidade Federal de Mato Grosso do Sul (UFMS/Campo Grande), onde atua como docente e pesquisadora na área da Fraseologia, investigando, principalmente, os seguintes temas: fraseologia e cultura, fraseo-

logia cognitivo-contrastiva, ensino e aprendizagem de línguas estrangeiras (espanhol). ORCID: <https://orcid.org/0000-0002-6308-9597>.

Fábio Henrique de Carvalho Bertonha

Doutor em Estudos Linguísticos pela Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP), campus de São José do Rio Preto (2022). Pós-doutorando pela Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP), campus de São José do Rio Preto (2023). Professor substituto da área de língua portuguesa e língua inglesa do IFSP, campus Barretos. Pesquisador na linha de Lexicologia e Lexicografia, da Linguística Aplicada. ORCID: <https://orcid.org/0000-0003-0770-4302>.

Guilherme Fromm

Doutor em Estudos Linguísticos e Literários em Inglês (USP), pós-doutorado em Linguística/Lexicografia (UFSCar). Professor Associado do ILEEL/UFU, área de Língua Inglesa na graduação e Léxico, na pós-graduação (PPGEL). Tem experiência na área de Linguística, com ênfase em Teoria e Análise Linguística, atuando principalmente nos seguintes temas: Lexicologia, Lexicografia, Terminologia, Terminografia bilíngue, Tradução, Linguística de Corpus e Análise/Treinamento/Desenvolvimento de software linguísticos. ORCID: <https://orcid.org/0000-0001-5654-0135>.

Ieda Maria Alves

Doutora em Linguistique pela Université Sorbonne-Nouvelle (Paris, França), Livre-docente pela Universidade de São Paulo. Realizou estágios de Pós-doutorado na Université Sorbonne Nouvelle (França), na Université Denis Diderot (França), na Université Paris 13 -

Villetaneuse (França), no Institut de la Langue Française (Nancy e Centre de Terminologie et de Néologie, Paris) e na Université Laval (Québec, Canadá). Professora titular sênior na Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo. Tem experiência na área de Linguística, com ênfase nos estudos do Léxico, atuando principalmente nos seguintes temas: Neologia, Morfologia, Lexicologia, Terminologia e Lexicografia. Uma das organizadoras da coleção *As Ciências do Léxico* (vol. III). ORCID: <https://orcid.org/0000-0002-1803-3615>.

Jessica dos Santos Paião

Doutora em Letras pelo Programa de Pós-Graduação em Letras da Universidade Federal de Mato Grosso do Sul – UFMS (2021). É docente da carreira do magistério do Ensino Básico, Técnico e Tecnológico (EBTT) no Instituto Federal de Mato Grosso do Sul - IFMS, campus Aquidauana. Atua na área do ensino de Língua Portuguesa e Língua Espanhola. Pesquisadora na área de Linguística, com ênfase nos seguintes temas: Fraseologia e Fraseografia das línguas portuguesa e espanhola. ORCID: <https://orcid.org/0000-0003-2771-265X>.

Márcia Sipavicius Seide

Doutora em Letras (Filologia e Língua Portuguesa) pela Universidade de São Paulo (2006) e Professora Associada da Universidade Estadual do Oeste do Paraná onde atua no ensino de graduação e de pós-graduação nas áreas de Onomástica, Lexicologia e Semântica. ORCID: <https://orcid.org/0000-0003-2859-1749>.

Márcio Sales Santiago

Doutor em Letras pela Universidade Federal do Rio Grande do Sul (UFRGS, 2013). Realizou estágio Pós-Doutoral em Terminologia na Universitat Pompeu Fabra (UPF, 2015), em Barcelona. Foi pesquisador do CNPq na Universidade Federal do Ceará (UFC, 2015), onde também realizou Pós-Doutorado em Linguística. Atualmente, é professor adjunto de Língua Portuguesa da Universidade Federal do Rio Grande do Norte (UFRN), professor permanente do Programa de Mestrado Profissional em Letras (Profletras), vice-diretor da Faculdade de Engenharia, Letras e Ciências Sociais do Seridó (FELCS) e líder do grupo de pesquisa Laboratório de Estudos Lexicais (LABLEX). Suas áreas de interesse são: Ciências do Léxico, Linguística Aplicada, Ensino de Língua Portuguesa e Tradução. ORCID: <https://orcid.org/0000-0001-9348-7273>.

Maria Cândida Trindade Costa de Seabra

Doutorado em Linguística pela mesma Instituição (2004). Desenvolve pesquisa em Linguística Teórica e Descritiva, nas perspectivas sincrônica e diacrônica, com ênfase em Onomástica (Toponímia e Antroponímia), Lexicologia, Lexicografia, Português do Brasil, Variação e Mudança Linguística. Foi coordenadora do GT de Lexicologia, Lexicografia e Terminologia da ANPOLL (2012-2014). É líder do Grupo Mineiro de Estudos do Léxico (GruMEL/UFMG/CNPq). Uma das organizadoras da coleção *As Ciências do Léxico* (vol. VI). ORCID: <http://orcid.org/0000-0003-4827-0635>.

Maria Celeste Augusto

Doutora em Semântica motivacional (1993) pela Universidade de Utreque, Países Baixos. Docente e investigadora (1982-2010) na Uni-

versidade de Utreque, Departamento Português / OTS. Desde 2010, Docente aposentada e Investigador visitante do Instituto de Linguística OTS, Utreque. Desde 1987, colaboradora do Atlas Linguistique Roman. Área de interesse mais recente: Paremiologia. ORCID: <http://orcid.org/0009-0002-5712-5544>.

Maria Cristina Parreira da Silva

Doutora em Linguística e Língua Portuguesa pela UNESP/Araraquara. Docente de FLE, no curso de Tradução, desde 2003 e no Programa de Pós-Graduação em Estudos Linguísticos, na linha de Lexicologia e Lexicografia, na UNESP/São José do Rio Preto. Orienta pesquisas em Estudos do Léxico e em Linguística Aplicada. Coautora do livro XERETANDO A LINGUAGEM: FRANCÊS. Membro do GTLEX da ANPOLL. Sócia do GEL-SP e da ABRALIN. Líder do GP GAMPLE (CNPq). ORCID: <https://orcid.org/0000-0002-8053-5413>.

Maria da Graça Krieger

Doutora em Linguística e Semiótica (USP, 1990). Pós-Doutorado em Terminologia: Universidad Pompeu Fabra, Barcelona (2002). Professora Titular aposentada de Língua Portuguesa da UFRGS. Fundadora do Projeto Termisul-UFRGS, atualmente, coordenadora emérita do Termisul. Pesquisadora em Lexicografia e Terminologia. Uma das organizadoras da coleção *As Ciências do Léxico* (vol. II). ORCID: <https://orcid.org/0000-0003-3159-0042>.

Maria José Bocorny Finatto

Linguista especializada em Estudos do Léxico e Terminologia. Investiga a história das terminologias médicas e os modos de aplicar técnicas da Linguagem Simples para promover maior acessibilidade

das informações escritas que trazem conteúdos de Utilidade Pública para diferentes perfis de pessoas no Brasil. É docente e orientadora de mestrado e doutorado junto à UFRGS. Coordena o grupo de pesquisa do DgP/CNPq “Acessibilidade Textual e Terminológica”. Uma das organizadoras da coleção *As Ciências do Léxico* (vol. IV). ORCID: <https://orcid.org/0000-0002-6022-8408>.

Martín Tapia Kwiecien

Magíster en Didáctica del Español como Lengua Extranjera (Universidad San Jorge Zaragoza- España). Actualmente, se desempeña como profesor titular de las cátedras “Fonética, Fonología, Morfología del Español y Principios de Contrastividad” y “Lexicología, Lexicografía Españolas y Principios de Contrastividad” ambas en las carreras de Español; y como profesor adjunto de la asignatura común “Lengua Castellana I” (FL-UNC). Escodirector de la Revista Digital de Lexicología, Lexicografía y Terminología (ReDILLeT). ORCID: <https://orcid.org/0009-0006-4142-3052>.

Odair Luiz Nadin

Doutor em Linguística e Língua Portuguesa pela Universidade Estadual Paulista “Júlio de Mesquita Filho” (Unesp - 2008). Livre-docente em Estudos do Léxico (Unesp-2018). Docente de Língua Espanhola no Departamento de Letras Modernas e na linha de Pesquisa Estudos do Léxico do Programa de Pós-Graduação em Linguística e Língua Portuguesa da Faculdade de Ciências e Letras da Unesp, campus de Araraquara. ORCID: <https://orcid.org/0000-0002-4655-0724>.

Pierre-André Buvet

Doutor em Sciences du Langage - Université Paris 13 (1993). Professor e pesquisador na Université Sorbonne Paris Nord, onde é diretor do Laboratório Théories Textes Numérique e responsável pelo Master Traitement Automatique des Langues/Traitement Informatique et Linguistique des Documents Écrits. Desenvolve pesquisas nos seguintes domínios da Linguística Teórica e Aplicada: Semântica, Léxico, Sintaxe, Diálogo Homem-Máquina, Inteligência Artificial, Lexicografia Eletrônica.

Renato Rodrigues-Pereira

Doutor em Linguística e Língua Portuguesa pela Universidade Estadual Paulista “Júlio de Mesquita Filho” – Unesp (2018). Docente da Universidade Federal de Mato Grosso do Sul – UFMS, nos cursos de: Licenciatura em Letras Português/Espanhol; Mestrado Profissional em Letras; Mestrado e Doutorado do Programa de Pós-Graduação em Letras (UFMS/CPTL); e Mestrado e Doutorado do Programa de Pós-Graduação em Estudos de Linguagens (UFMS/FAALC). Pesquisador na área de Linguística, com ênfase em Lexicografia Geral, Lexicografia Pedagógica, Toponímia e Linguística Aplicada ao Ensino de Línguas. ORCID: <https://orcid.org/0000-0001-9870-3780>.

Sabrina Pereira de Abreu

Doutora em Letras pela Pontifícia Universidade Católica do Rio Grande do Sul – PUCRS (1998). Professora titular aposentada pela Universidade Federal do Rio Grande do Sul (UFRGS). Docente-convidada junto ao Programa de Pós-Graduação em Letras da UFRGS. Pesquisadora na área de Estudos do Léxico: aspectos morfológicos, sintáticos e semânticos. Realizou estágio sênior no Laboratoire

de Analyse et Traitement Informatique de la Langue Française da Universidade de Lorraine, França (2015-2016). ORCID: <https://orcid.org/0000-0001-7657-3655>.

Sara Augusto Carra

Mestra em Estudos da Linguagem pela Universidade Federal do Rio Grande do Sul (UFRGS) desde 2019. Em dezembro de 2020, ingressou no doutorado em Estudos da Linguagem na mesma universidade, onde também participa do projeto Expressões fraseológicas: um estudo teórico-prático segundo o ponto de vista da TST, coordenado pela professora doutora Sabrina Pereira de Abreu. Sua pesquisa é voltada diretamente a frasemas e dicionários. ORCID: <https://orcid.org/0000-0002-3735-123X>.

‡

Este livro foi editorado com as fontes Crimson Text e Montserrat.
Publicado on-line em: <https://repositorio.ufms.br>

hum pão de assucar por cada quinze; e isto pela renda da terra [...]. LUIZ DOS SANTOS VILHENA (1921) [1802] CARTA QUINTA [A00_0407 p. 182]. 1ª. datação [1552] Ó trato bendito, não de açuquere corruptível, mas de graça mais saborosa que favo de mel, quam poucos há que te queirão ter! P. MANUEL DA NÓBREGA (1956) [1552] CARTA DO P. MANUEL DA NÓBREGA [AOS MORA DORES DE PERNAMBUCO], BAÍA [5 DE JUNHO 1552 [A00_0007 p. 334].

MEL DE PAU

Mel das abelhas uruçú, jataí e outras que o ajuntam em ocos das árvores. [...] então fe lhe lâcem duas oytavas de fenne, com que dê huma leve fervura, e tirado do fogo fe coe, e torne a elle a purificar com fete, ou oyto onças de mel de pão, dando algumas fervuras, e sendo neccffario coado, fe coe, e guarde para o ufo. LUIS GOMES FERREIRA (1735) [1735], DAS OBÍTRUCÇOENS [B00_0030 p. 82].

MEL DE ENGENHO

Caldo da cana cozido e grosso que se apura para ir às formas a purgar. MEL DE ENGENHO Ouve-se falar n'ell' tão a miudo que entender-se-á talvez, que esta é uma das grandes manufacturas. Examinando o caso está visto, que as maiores quantidades não passam de algumas frasqueira d'elle, que os fabricantes vendem á razão de 320 rs. cada frasco para supprir a falta do assucar. ALEXANDRE RODRIGUES FERREIRA [séc. XVIII], 2.ª PARTE: BAIXO RIO NEGRO - PARTICIPAÇÃO SÉTIMA: PARTICIPAÇÃO GERAL DO RIO NEGRO [A00_2243 p. 673].

VINHO DE MEL

Bebida alcoólica feita a partir da fermentação de caldas de açúcar de cana ou de frutas; também denominada cachaca [...] e com tudo a experiencia mostrou que este proveito particular era de muito pouca consideração a respeito do damno publico que sepadecia com as vendas do dito **vinho demel**, acujas cazas acodião de ordinario denoite e de dia os escravos com os furtos que fazião aseos Senhores, [...]. RUI CARVALHO PINHEIRO (1944) [1635], ESCRITURA QUE SE FEZ SOBRE A PROHIBIÇAM DO VINHO DE MEL [A00_1374 p. 281]. 1ª. datação [1500] [...] de ranlhes aly de comer pam e pescado cozido, confeitos fartees mel e figos pasados. nõ quiseram comer daquily casy nada e algu a coussa se aprouauam lamcauãna logy fora. trouueranlhes vinho perh a taça. PERO VAZ DE CAMINHA (1964) [1500], CARTA DE PERO VAZ DE CAMINHA [A00_0335 p. 3].

Aparecida Negri Isquierdo

Professora nos Programas de Pós-Graduação em Estudos de Linguagens – Campo Grande e em Letras – Três Lagoas, da Universidade Federal de Mato Grosso do Sul.

Coordenadora do Projeto DTMS – Dicionário de Topônimos de Mato Grosso do Sul.

Membro do Comitê Nacional do Projeto ALiB – Atlas Linguístico do Brasil, como Diretora Científica e Coordenadora do Projeto ALiB – Regional Mato Grosso do Sul.

Vice-coordenadora do GT de Lexicologia, Lexicografia e Terminologia da ANPOLL, biênios 1998-2000 e 2004-2006.

Coordenadora do GT de Lexicologia, Lexicografia e Terminologia da ANPOLL, biênio 2006-2008.

Bolsista Produtividade em Pesquisa do CNPq.

Elizabeth Aparecida Marques

Professora nos Programas de Pós-Graduação em Estudos de Linguagens – Campo Grande e em Letras – Três Lagoas, da Universidade Federal de Mato Grosso do Sul.

Coordenadora do Projeto Tesouro do Léxico Patrimonial Galego e Português: Foco sobre a Região Centro-Oeste do Brasil.

Coordenadora do Projeto Variação, recursos, territórios: fraseologia e tratamento informatizado.

Chefe da Secretaria da Editora UFMS – SEDIT/AGECOM.

Vice-coordenadora do GT de Lexicologia, Lexicografia e Terminologia da ANPOLL, entre 2017 a 2021.

Coordenadora do GT de Lexicologia, Lexicografia e Terminologia da ANPOLL, biênio 2021-2023.

A coleção *As Ciências do Léxico* chega ao seu décimo volume mantendo-se fiel aos objetivos e estrutura da primeira edição, em 1998, à época definidos pela sua idealizadora, Maria Tereza Camargo Biderman, em função do propósito da publicação. Desta forma, os trabalhos que compõem este volume, obedecendo ao plano da obra, estão distribuídos segundo uma das três grandes partes em torno das quais os textos são organizados: Parte I – Lexicologia; Parte II – Lexicografia e, Parte III – Terminologia.

O volume X da obra *As Ciências do Léxico*, como os demais da coleção, publica artigos que resultam de investigações desenvolvidas por pesquisadores vinculados ao Grupo de Trabalho em Lexicologia, Lexicografia e Terminologia da ANPOLL – Associação Nacional de Pós-Graduação em Letras e Linguística, trabalhos discutidos nos encontros promovidos pelo grupo. Neste volume a obra abriga também artigos produzidos por dois pesquisadores estrangeiros, um francês e uma portuguesa.

A exemplo dos anteriores, o volume X tem como público-alvo mais amplo pesquisadores e estudantes de pós-graduação em Linguística e Língua Portuguesa interessados em estudos sobre o léxico em suas múltiplas abordagens e, de forma mais pontual, interessa a docentes que atuam na Educação Básica à medida que também reúne capítulos que tratam da questão do ensino do léxico e do uso do dicionário como recurso pedagógico no ensino da língua materna e de línguas estrangeiras.



NEGRO - PARTICIPAÇÃO SÉTIMA: PARTICIPAÇÃO GERAL DO RIO NEGRO [A00_2243 p. 673].

VINHO DE MEL

Bebida alcoólica feita a partir da fermentação de caldas de açúcar de cana ou de frutas; também denominada cachça. [...] e com tudo a experiencia mostrou que este proveito particular era de muito pouca consideração a respeito do damno publico que sepadecia com as vendas do dito **vinho demel**, acujas cazas acodião de ordinario denoite e de dia os escravos com os furtos que fazião aseos Senhores, [...]. RUI CARVALHO PINHEIRO (1944) [1635], ESCRIPTURA QUE SE FEZ SOBRE A PROHIBIÇAM DO VINHO DE MEL [A00_1374 p. 281]. 1ª. datação [1500] [...] de ranlhes aly de comer pam e pescado cozido. confejtos fartees mel e figos pasados. nõ quiseram comer daquilo casy nada e algu a coussa se aprouauam lamçauãna logo fora. trouueranlhes vinho perh a taça. PERO VAZ DE CAMINHA (1964) [1500], CARTA DE PERO VAZ DE CAMINHA [A00_0335 p. 3].