

Plato

REVISTA DO INSTITUTO INTERNACIONAL
DA LÍNGUA PORTUGUESA

NÚMERO
09
VOLUME 5
2022

**NÍVEL LEXICAL: AS PALAVRAS E OS
TERMOS. DESAFIOS POLÍTICO-
LINGUÍSTICOS PARA O PORTUGUÊS,
LÍNGUA PLURICÊNTRICA.**

IILP e Itamaraty (Organizadores)



Contato:

Instituto Internacional da Língua Portuguesa (IILP)
Avenida Andrade Corvo, nº 8, CP 382
Plateau, cidade da Praia, Cabo Verde
Telefone: +238 261 95 04
<https://iilp.cplp.org/plato/>
E-mail: platorevista@gmail.com

**Nível lexical: as palavras e os termos.
Desafios político-linguísticos para o português, língua pluricêntrica.**

Número 9

Volume 5

2022

Organizadores

Instituto Internacional da Língua Portuguesa (IILP)
Ministério das Relações Exteriores do Brasil – Itamaraty

Editor da Platô

Instituto Internacional da Língua Portuguesa

Diretor Executivo do IILP

Incanha Intumbo

Comité Editorial Platô

Ana Isabel Soares (Instituto Camões, Portugal)
António Branco (Universidade de Lisboa, Portugal)
Emir José Suaiden (Instituto Brasileiro de Informação em Ciência e Tecnologia/IBICT, Brasil)
Gregório Firmino (Universidade Eduardo Mondlane, Moçambique)
Guadelupe Teresinha Bertussi (Universidade Pedagógica do México, México)
Manuel Monteiro da Veiga (Universidade de Cabo Verde, Cabo Verde)
Maria José Grosso (Universidade de Macau, Macau)

Equipa Técnica de Editoração Platô

Rosângela Morello: supervisão técnica
Isabelle Garcia Morello: design e diagramação

Criação da Capa Platô

Isabelle Garcia Morello

**Platô Revista do Instituto Internacional da Língua Portuguesa - Instituto
Internacional da Língua Portuguesa (IILP)**

N.9, V.5 (2022), Cidade da Praia, Cabo Verde
Editora do IILP, 2022.
Semestral

ISSN: 2311-6625 on-line

1. Língua Portuguesa - Periódicos. 1. Instituto Internacional da Língua Portuguesa
Todos os direitos autorais estão reservados a PLATÔ/IILP

ÍNDICE

04	Apresentação - Instituto Internacional da Língua Portuguesa	Incanha Intumbo
06	Apresentação - Ministério das Relações Exteriores do Brasil	Instituto Guimarães Rosa
08	Apresentação - Coordenação do Projeto Terminologias Científicas e Técnicas Comuns da Língua Portuguesa	Gladis Almeida
10	Projetos pluricêntricos do IILP. Apontamentos para a história da política linguística do português	Margarita Correia
28	Variedades nacionais de línguas pluricêntricas: o caso do português em Moçambique	Inês Machungo e Gregório Firmino
46	A importância das terminologias em português para a tradução e a interpretação na União Europeia	Joana Guimarães
56	A língua portuguesa no espaço dos países africanos de língua oficial portuguesa - da variedade do português de São Tomé e Príncipe à definição de um padrão	Abigail Tiny Cosme
68	A língua portuguesa no espaço dos países africanos de língua oficial portuguesa - uma reflexão sobre Cabo Verde	Adelaide Monteiro
76	Construindo uma terminologia de raiz: textos legislativos sob exploração terminológica	Maria José Bocorny Finatto, Francine Facchin Esteves e Guillermo Silva Villar
98	Língua e diplomacia - uma agenda para o diálogo em língua portuguesa	Margarita Correia
116	Procedimentos teórico-metodológicos para recolha e sistematização de terminologias no âmbito da Comunidade de Países de Língua Portuguesa (CPLP)	Gladis Barcellos Almeida e Tanara Zingano Kuhn

Apresentação

Instituto Internacional da Língua Portuguesa

Incanha Intumbo (Diretor Executivo do IILP)

O Instituto Internacional da Língua Portuguesa (IILP), instituição da Comunidade de Países da Língua Portuguesa (CPLP) cuja missão é a promoção, a difusão e a internacionalização da Língua Portuguesa (LP) lança, com a colaboração da Embaixada de Brasil na Praia, mais uma edição da revista Platô, o volume 5, número 9, que contempla comunicações proferidas no âmbito da mesa-redonda **A língua portuguesa no espaço dos Países Africanos de Língua Oficial Portuguesa (PALOP)**, realizada por ocasião do II Formação das Terminologias Científicas e Técnicas Comuns (TCTC) realizada em Brasília, novembro de 2021, promovida pelo IILP. No evento, participantes dos diferentes países da CPLP apresentaram comunicações que versaram sobre questões da língua portuguesa, as complexas realidades sociolinguísticas de alguns dos EM dos PALOP, questões do léxico e da terminologia. Este número especial da Platô reúne os textos de alguns dos participantes.

As TCTC é um projeto do IILP que visa a criação e a disponibilização pública, tanto quanto possível, de recursos terminológicos técnicos e científicos para os Estados-Membros da CPLP, integrado no portal do Vocabulário Ortográfico Comum (VOC) da LP. Pela sua riqueza, pertinência e atualidade, a divulgação dos textos apresentados pelos conferencistas tornou-se uma prioridade para o IILP, mas também uma oportunidade para brindar os nossos habituais leitores, especialistas ou simples interessados em LP, com os conteúdos dos mesmos.

Os artigos abordam diferentes temas sobre a LP, incluindo a importância das terminologias em português para a tradução e a interpretação na União Europeia, orado pela Doutora Joana Guimarães, professora auxiliar da Faculdade de Letras da Universidade do Porto, que realça a importância de um sólido trabalho de terminologia, desenvolvido por terminólogos em cooperação com tradutores e intérpretes e com os especialistas dos mais variados setores profissionais.

Adelaide Monteiro, do Instituto do Património Cultural, órgão do Ministério da Cultura e das Indústrias Criativas de Cabo Verde, descreve os usos da LP em Cabo Verde, onde a interação da LP com as línguas Oeste Atlânticas sobre tudo as da Costa Ocidental da África, de onde eram originários os primeiros habitantes africanos que contribuíram no povoamento de CV e onde a LP se hibridizou com as mencionadas línguas africanas, dando génese ao crioulo de base portuguesa falado em CV, o cabo-verdiano.

Abigail Tiny Cosme, docente no Departamento de Língua Portuguesa da Universidade de São Tomé e Príncipe (STP), descreve a situação sociolinguística de STP, desde o início do povoamento das ilhas por povos de origem bantu, primeiro São Tomé no século XV depois Príncipe no início do século XVI. O texto aborda também o nascimento de um pidjin, apontando o período a seguir a 1520 como o período de rápido declínio do comércio do delta do Níger. E de contactos entre portugueses e os escravizados resgatados do delta do

Níger, sobretudo do antigo Reino de Benim, resultou a criação de condições ideais para a formação do pidgin que daria origem depois ao proto-crioulo do Golfo da Guiné (PCGG) que se terá ramificado, dando origem aos quatro CGG: Forro, Angolar, Lung'le e Fa d'Ambô.

Margarita Correia, da Universidade de Lisboa e CELGA-ILTEC e Presidente do Conselho Científico do IILP (2018-2022), faz um depoimento, resultado de uma reflexão sobre a sua visão do que é e do que poderia ser a função dos diplomatas na promoção da língua portuguesa e do que poderiam ser alguns aspetos da 'diplomacia linguística' relacionados com o português nestas instituições e nos países nelas representados e termina tecendo considerações sobre alguns tópicos para a construção de uma agenda que promova o diálogo em língua portuguesa. O segundo texto da Professora Doutora Margarita Correia traça o percurso histórico dos projetos pluricêntricos do IILP, sublinhando a nova era de gestão da LP caracterizada pelo multilateralismo, impulsionada pelos DE Gilvan Müller de Oliveira (2010-2014) e Marisa Mendonça (2014-2018).

A equipa constituída por Maria José Bocorny Finatto, Francine Facchin Esteves e Guillermo Silva Villar, docente, mestranda e graduando respetivamente, todos da UFRGS, no texto sobre Construindo uma terminologia de raiz: textos legislativos sob exploração terminológica, sintetiza os princípios e procedimentos básicos para quem deseja dar os seus primeiros passos no trabalho de reconhecimento terminológico (RT) baseado em fontes escritas (o corpus) em formato digital.

Gladis Maria de Barcellos Almeida, docente na Universidade Federal de São Carlos e Tanara Zingano Kuhn, investigadora do Centro de Estudos de Linguística Geral e Aplicada da Universidade de Coimbra – CELGA-ILTEC, contextualizam o projeto das TCTC, abordando a sua origem e relevância.

Inês Machungo e Gregório Firmino, da Universidade Eduardo Mondlane analisam a variedade do Português de Moçambique, o primeiro país a produzir um dicionário de Português fora de Brasil e Portugal.

Esta soberba lista de temas analisados por diferentes personalidades de diferentes instituições académicas e de investigação científica da nossa comunidade traz à luz diferentes aspetos da LP, reveladoras das dinâmicas resultantes das interações da mesma com as realidades linguísticas e sociais nacionais dos EM, provando que o contacto linguístico tem promovido um maior enriquecimento da LP, que a variação é uma realidade na LP e que o compromisso de promoção e difusão da nossa língua, nas suas diferentes realizações, deve ser um compromisso de todos nós.

Votos de boa leitura.

Apresentação

Ministério das Relações Exteriores do Brasil

Instituto Guimarães Rosa

Dando continuidade a uma trajetória bem-sucedida, a revista Platô, publicada pelo IILP, consolida-se a cada edição como periódico de referência no plano das ciências humanas em língua portuguesa. Nessa tradição inovadora, a revista trata, no presente número, de tema de suma importância à gestão multilateral de nosso idioma: as Terminologias Científicas e Técnicas Comuns da Língua Portuguesa (TCTC).

Esta edição da Platô é fruto de longa série de debates, ações, de políticas levadas adiante no âmbito da CPLP acerca das TCTC. Ao longo de seus artigos, sintetiza e aprofunda intervenções de alto nível, ocorridas durante a 2ª Formação Técnica do projeto TCTC, entre 22 e 25 de novembro de 2021, na Universidade de Brasília.

Organizada pela Direção Executiva do IILP e contando com apoio e participação do Instituto Guimarães Rosa (IGR), a formação técnica decorreu de proposta e compromissos financeiros assumidos pelo Brasil, por ocasião da IV Reunião Extraordinária do Conselho Científico do IILP (18 a 20 de novembro de 2020) e da XVII Reunião Ordinária (14 e 15 de junho de 2021).

As oficinas da formação técnica foram ministradas pelas professoras Gladis Barcellos Almeida e Tanara Kuhn, da Equipe Central do projeto TCTC, sendo extremamente produtivas. Trataram da história do projeto TCTC, sua motivação original, objetivos, desafios, cronograma e resultados esperados, bem como de possíveis abordagens teóricas no campo das terminologias. Ao longo das oficinas, a equipe central orientou-se por metodologia de trabalho ancorada no pluricentrismo, bem como nas diferenças culturais entre as diversas variantes da língua portuguesa.

Nesse contexto, também realizou-se exercício de harmonização do léxico do ensino superior dos países da CPLP, com vistas a contribuir para maior mobilidade acadêmica no seio da comunidade.

A reação positiva dos presentes foi atestada pelo alto nível das intervenções, reunidas neste número especial da Platô que o leitor tem em mãos. Abordando temas amplos, inter-relacionados, trazendo contribuições de pesquisadores de ponta de países de língua portuguesa, os artigos aqui presentes certamente contribuirão para o avanço do debate no campo das TCTC e para a construção de um repertório terminológico comum.

Boa leitura!

Embaixadora Paula Alves de Souza

Diretora do Instituto Guimarães Rosa (IGR)

Apresentação

Coordenação do Projeto Terminologias Científicas e Técnicas Comuns da Língua Portuguesa

Gladis Barcellos Almeida

Em novembro de 2021, realizou-se, em Brasília (Brasil), o segundo curso de formação para o desenvolvimento do projeto Terminologias Científicas e Técnicas Comuns (TCTC) no espaço da Comunidade dos Países de Língua Portuguesa (CPLP).

Para além da formação em si, na ocasião, colegas dos diferentes países da CPLP realizaram importantes intervenções, dando voz a questões da língua portuguesa, como também a temas atinentes ao léxico em geral e à Terminologia em particular.

Todos os participantes foram convidados a enviar os textos das suas intervenções. Assim, neste número especial da Revista Platô, estão reunidos os textos dos autores que aceitaram nosso convite.

Seguindo a ordem em que foram apresentadas no evento supracitado, este número se inicia com o artigo de Margarita Correia (Portugal), intitulado “Projetos pluricêntricos do IILP. Apontamentos para a história da política linguística do português,” no qual a autora destaca os projetos gerenciados pelo Instituto Internacional da Língua Portuguesa (IILP) desde 2010, chamando atenção para “uma nova era de gestão das políticas de língua portuguesa,” bem como para as condições e fatos que originaram esse ciclo de projetos pluricêntricos. No texto, a autora também explora os conceitos de língua pluricêntrica, pluricentricidade e pluricentrismo, de modo a delimitá-los e defini-los.

Na sequência, temos o artigo de Inês Machungo e Gregório Firmino (Moçambique), intitulado “Variedades nacionais de línguas pluricêntricas: o caso do português em Moçambique”, em que os autores abordam questões atinentes às variedades não-nativas (VNN) de línguas coloniais, as quais “expandiram o seu espaço de uso e papel social como entidades simbólicas e meios de comunicação comunitária” e vêm se nativizando. Os autores tratam especificamente do português de/em Moçambique, apresentando-nos os desafios que se lhe colocam bem como sua dimensão simbólica e linguística. Para ilustrar esse processo de nativização do português, os autores elegem o nível lexical.

Em seguida, Joana Guimarães (Portugal) apresenta seu artigo “A importância das terminologias em português para a tradução e a interpretação na União Europeia”. Nesse texto, a autora discorre sobre os desafios enfrentados pelos tradutores/intérpretes naquele cenário multilíngue. Apesar de referir uma série de recursos e ferramentas que auxiliam, em certa medida, o trabalho desses profissionais, a autora chama a atenção para o fato de estarem dispersos e sem alinhamento entre eles. No que concerne à língua portuguesa, a autora ressalta que “ela é na maioria das vezes uma língua ‘não original’, ou seja, aquilo a que no âmbito dos Estudos de Tradução se chama uma língua de chegada, uma língua para a qual se traduz.” O que significa dizer que uma grande quantidade de termos entra constantemente na variedade europeia do português. Dessa forma, o português vai sendo integrado, paulatinamente, às memórias de tradução, bases de dados terminológicas ou memórias

documentais multilíngues, fato que evidencia a necessidade de se produzirem terminologias de qualidade nesta língua.

Deparamo-nos, na sequência, com o artigo “A língua portuguesa no espaço dos países africanos de língua oficial portuguesa: da variedade do português de São Tomé e Príncipe à definição de um padrão,” de Abigail Tiny Cosme (São Tomé e Príncipe). No texto, a autora nos apresenta o panorama linguístico do país, considerando os eventos históricos e políticos que colaboraram para uma grande diversidade linguística e cultural. A despeito de tratar-se de um país pequeno, com cerca de 200 mil habitantes, “um conjunto de línguas compõe o mosaico linguístico do país:” a língua portuguesa, o forro, o lung’ie, o angolar, o crioulo de Cabo Verde, o português dos tongas e resquícios de línguas do grupo bantu. Considerando esse cenário, a autora analisa a situação linguística, apontando especialmente “as idiossincrasias do Português de e em São Tomé e Príncipe, fundamentais para um olhar crítico sobre o surgimento de uma variedade nacional, bem como da respetiva norma-padrão.”

Temos, a seguir, o artigo intitulado “A língua portuguesa no espaço dos países africanos de língua oficial portuguesa – uma reflexão sobre Cabo Verde,” de Adelaide Monteiro (Cabo Verde). No texto, a autora discorre sobre o uso da língua portuguesa em Cabo Verde, que convive desde meados do século XVI com a língua nativa, a língua cabo-verdiana. A autora chama atenção para um conjunto significativo de itens léxicos comuns entre as duas línguas, apesar de haver um distanciamento no que se refere aos níveis sintático e morfológico. Com relação à “prática linguística do cabo-verdiano no seu dia a dia,” ela é definida por ambas as línguas, já que o que determina a escolha é o contexto (formal e informal) e a modalidade (oral e escrita). Nesse artigo, portanto, a autora faz uma “abordagem de gestão de línguas em situação de contato linguístico,” oferecendo-nos um panorama sociolinguístico de Cabo Verde.

Logo em seguida, temos o artigo “Construindo uma terminologia de raiz: textos legislativos sob exploração terminológica”, de autoria de Maria José Bocorny Finatto, Francine Facchin Esteves e Guillermo Silva Villar, que retrata a experiência vivida pelos participantes na oficina 1 do curso em Brasília, na qual se propunha uma metodologia para a criação de uma terminologia a partir do zero. Nesse artigo, os autores apresentam uma síntese dos princípios básicos que norteiam a tarefa bem como os procedimentos para a identificação de candidatos a termos a partir de corpus digital. Para tanto, utilizam-se de textos legislativos do Brasil como corpus e da ferramenta computacional AntConc, especialmente as funcionalidades “wordlist (lista de palavras); concordance (gerador de contextos por palavra ou expressão de busca); e n-grams ou clusters (grupos de palavras que se repetem ao longo de um texto ou arquivo).”

Ainda de acordo com a cronologia do evento, temos o artigo “Língua e diplomacia – uma agenda para o diálogo em língua portuguesa”, de autoria de Margarita Correia (Portugal). Nesse texto, a autora faz depoimentos pessoais e propõe reflexões por meio das quais expõe a sua “visão do que é e do que poderia ser a função dos diplomatas na promoção da língua portuguesa.”

Por fim, Gladis Barcellos Almeida e Tanara Zingano Kuhn (ambas do Brasil) apresentam em seu artigo os “Procedimentos teórico-metodológicos para recolha e sistematização de terminologias no âmbito da Comunidade de Países de Língua Portuguesa (CPLP).” O texto procura contextualizar o projeto Terminologias Científicas e Técnicas Comuns (TCTC), detalhando sua origem e relevância para a CPLP, bem como seu embasamento teórico e etapas metodológicas.

Agradeço profundamente aos autores aqui reunidos, que aceitaram o convite de publicação e tornaram este número tão rico.

Aos leitores, desejo que desfrutem!

Projetos pluricêntricos do IILP: apontamentos para a história da política linguística do português¹

Margarita Correia

Sobre o autor

Universidade de Lisboa / CELGA-ILTEC. Presidente do Conselho Científico do IILP entre 2018 e 2022.

Resumo

Conclui-se no final deste ano de 2022 um impressionante ciclo de projetos pluricêntricos do Instituto Internacional da Língua Portuguesa (IILP), ciclo que teve início em 2010. Além de terem trazido renovadas energias à Instituição, estes projetos permitiram inaugurar uma nova era de gestão das políticas de língua portuguesa. Importa registar para memória futura as condições e factos que se encontram na génese deste fenómeno, bem como os principais marcos desta jornada, que mudou a conceção de gestão da política linguística do português e os modos de pensar multilateralmente a língua, que constitui o cimento fundador da própria Comunidade de Países de Língua Portuguesa (CPLP). O registo destes factos constitui o principal objetivo deste trabalho.

Para atingir esse objetivo, o presente texto encontra-se dividido em três partes principais. Em primeiro lugar, serão abordados os conceitos de língua pluricêntrica, pluricentricidade e pluricentrismo, procurando-se defini-los e delimitá-los. Na segunda parte, serão apresentados dois factos fundamentais para o impulso inicial deste movimento: a aplicação do Acordo Ortográfico da Língua Portuguesa, de 1990, e a elaboração do Plano de Ação de Brasília na sequência da realização da I Conferência Internacional sobre o Futuro da Língua Portuguesa no Sistema Mundial, em 2010. Em seguida, providenciar-se-á uma breve panorâmica sobre o nascimento do Vocabulário Ortográfico Comum da Língua Portuguesa (VOC), primeiro grande projeto pluricêntrico do IILP, dando conta das estreitas relações que este estabeleceu com outros projetos pluricêntricos, a saber, o Portal do Professor de Português Língua Estrangeira-Língua Segunda (PPPLE) e as Terminologias Científicas e Técnicas Comuns da Língua Portuguesa (TCTC). O trabalho apresenta algumas considerações finais.

¹ Este texto decorre da preparação da conferência de abertura do 2.º Curso de Formação para o desenvolvimento do projeto das TCTC no espaço da CPLP (Brasília, outubro de 2021).

Palavras-chaves

IILP, política linguística, língua pluricêntrica, pluricentrismo, VOC, TCTC, PPPL

1. Introdução

O Instituto Internacional da Língua Portuguesa (IILP), criado em 1989, ainda antes da própria Comunidade de Países de Língua Portuguesa (CPLP), encontra-se a ponto de concluir um impressionante período de 13 anos de intensa produtividade, que se traduziu na conceção e realização de projetos de grande fôlego, de matriz pluricêntrica, que, além de terem trazido renovadas energias à Instituição, permitiram inaugurar uma nova era de gestão partilhada das políticas linguísticas para a língua portuguesa. Estes projetos deram corpo às atribuições cometidas ao IILP pelos seus Estatutos, aprovados em 2010 em Conselho Científico e pelo XV Reunião Ordinária do Conselho de Ministros da Comunidade de Países de Língua Oficial Portuguesa, realizada em Luanda em 22 de julho de 2010, mas que até hoje só Portugal ratificou². Até então, a sua atividade era praticamente inexistente ou desconhecida, facto referido em literatura da época (e.g. Faraco, 2012; Neto, 2013). Com efeito, nos últimos 13 anos, o IILP, no qual se encontram representados os nove Estados-membros da CPLP, veio a constituir-se como entidade supranacional de gestão internacional da língua portuguesa, tendo contribuído para o enfraquecimento de nacionalismos linguísticos, isolacionistas e anacrónicos, que ainda vão perdurando nos dias que correm.

Infelizmente, o trabalho desenvolvido no IILP, pelos seus diretores executivos³, funcionários⁴ e por todos aqueles que têm trabalhado nos bastidores, está longe de receber o (re)conhecimento devido das sociedades, dos governos e das próprias estruturas da CPLP. Parece que estas entidades ainda não terão compreendido o enorme potencial de desenvolvimento dos países da CPLP que a língua portuguesa e o IILP encerram, o que constitui um enorme desperdício de oportunidades efetivas de fortalecimento do português e dos povos da CPLP⁶.

Com este trabalho, pretendemos contribuir para contrariar o esquecimento e o descaso a que o IILP tem sido votado, sendo este o seu primeiro objetivo.

Como escreveu Camões, “Todo o mundo é composto de mudança, / Tomando sempre novas qualidades”. Esta é, assim, tão inevitável quanto saudável. Estamos num momento de passagem de testemunho a novas gerações. Importa registar os principais marcos desta jornada, para que as novas gerações possam, se assim o entenderem, aprender com as vitórias e derrotas do passado. Este constitui, então, o segundo objetivo deste texto.

² De acordo com o artigo 4.º, “2 — São atribuições gerais do IILP: (...)

d) Promover e acompanhar o desenvolvimento de instrumentos de normalização linguística comum;

e) Propor e ou avaliar projetos multilaterais de apoio ao desenvolvimento e otimização das competências em Língua Portuguesa em Estados membros da CPLP; (...).”

³ Destaque-se o exercício das funções de Diretor Executivo de Gilvan Müller de Oliveira (2010- 2014) e Marisa Mendonça (2014-2018).

⁴ Destaque-se o incomensurável contributo dado à instituição, em todos estes anos e especialmente desde 2019, por Maria José Sousa, Denise Fonseca e Nadine Rodrigues.

Para atingir estes dois objetivos, o trabalho encontra-se dividido em três partes principais. Após esta introdução, partindo do caso da língua portuguesa, procurar-se-á definir e delimitar os conceitos de 'língua pluricêntrica', 'pluricentricidade' e 'pluricentrismo', explicando como eles entraram no vocabulário do IILP, da CPLP e dos governos de alguns Estados-membros. Na terceira parte, serão apresentados dois factos que consideramos basilares para o impulso inicial da trajetória percorrida pelo IILP: a aplicação do Acordo Ortográfico da Língua Portuguesa de 1990 (AOLP90) e a elaboração do Plano de Ação de Brasília para a Promoção, a Difusão e a Projeção da Língua Portuguesa (PAB2010), produto da I Conferência Internacional sobre o Futuro da Língua Portuguesa no Sistema Mundial. Em seguida, providenciar-se-á uma breve panorâmica histórica sobre a génese dos projetos pluricêntricos do IILP, com ênfase no VOC; será feita referência ao Portal do Professor de Português Língua Estrangeira-Língua Segunda (PPPLE) e às Terminologias Científicas e Técnicas Comuns da Língua Portuguesa (TCTC). Concluimos o trabalho com algumas considerações finais.

2. O português, a pluricentricidade e o pluricentrismo

Muhr (2016), tomando por base a proposta de Clyne (1992), especifica que uma língua pluricêntrica é uma língua que é usada em pelo menos duas nações, nas quais tem estatuto oficial, como língua do estado, colíngua de estado ou língua regional, com normas (codificadas), que normalmente contribuem para o reforço da identidade nacional ou pessoal, tornando cada nação um centro produtor de norma, pelo uso deliberado das suas normas específicas.

A situação do português no início do século XXI poderia ser descrita como existindo duas variedades-padrão nacionais estabilizadas, com normas codificadas, a do Brasil e a de Portugal, que constituem as variedades dominantes do português; a par destas existem seis outras potenciais variedades nacionais (Angola, Cabo Verde, Guiné-Bissau, Moçambique, São Tomé e Príncipe Timor-Leste), que apresentam diferentes graus de implantação entre a população, reconhecimento e descrição, destacando-se, a este respeito, Moçambique^{7, 8}. Ao longo deste século, a situação linguística de vários países de língua oficial portuguesa tem evoluído. Angola e Moçambique, em particular, têm vivido um crescimento exponencial em termos geostratégicos, económicos, sociais e demográficos, muito motivado pelo final das sangrentas guerras civis que se seguiram à independência; não é de estranhar que tenham

⁵ Diversos artigos de opinião publicados nos últimos tempos têm apontado falhas de funcionamento e de definição de políticas por parte da CPLP – cf. já em 2022, os de João Melo (<https://www.dn.pt/opiniao/uma-comunidade-que-nao-se-conhece-nao-e-uma-comunidade-15063163.html>) e Rosália Amorim (<https://www.dn.pt/edicao-do-dia/18-jul-2022/a-lusofonia-e-as-suas-oportunidades-15027067.html>).

⁶ "A pluricentric language is a language that is used in at least two nations where it has an official status as state language, co-state language, or regional language with its own (codified) norms that usually contribute to the national/personal identity, making the nation a norm-setting centre by the deliberate use of the norms native to this specific nation."

⁷ Além dos projetos mencionados neste artigo, obras como a de Gonçalves (2010) constituem a afirmação clara da moçambicanidade do português.

⁸ Cf. Pinto (2021) e Pinto (no prelo), para a caracterização do pluricentrismo da língua portuguesa. Cf., ainda, o recém-publicado n.º 42 (2022) da revista Seminal (UERJ), subordinado ao tema "Português, língua de muitas pátrias." - <https://www.e-publicacoes.uerj.br/index.php/cadernoseminale/issue/view/2672/showToc>

também assistido a um crescimento exponencial do uso português não apenas como língua segunda, mas também como língua materna, muito devido à expansão do sistema educativo, especialmente nos grandes centros urbanos e entre os mais jovens. Os dois países apresentam, no entanto, posturas bem diferenciadas relativamente à sua variedade nacional do português. Em Moçambique, assiste-se ao reconhecimento coletivo da existência de uma variedade moçambicana da língua portuguesa, fator relevante para a definição da ‘moçambicanidade’, e ao desenvolvimento sistemático de trabalhos descritivos e de importantes instrumentos de codificação dessa variedade nacional, tais como o Vocabulário Ortográfico do Português de Moçambique (Machungo, org. 2017) e o Dicionário do Português de Moçambique (DiPoMo), projeto em curso. Ao contrário, em Angola, assiste-se a um discurso oficial que rejeita o inevitável desenvolvimento de uma variedade nacional própria, não a reconhecendo e reportando-se sistematicamente ao português europeu como variedade-padrão, o que se traduz na escassez de trabalhos descritivos sistemáticos, na ausência de qualquer ímpeto codificador do português de Angola e em claros prejuízos ao nível do ensino.

A situação dos restantes países de língua portuguesa é menos propícia à emergência e estabilização de variedades nacionais próprias. Um caso a destacar é, no entanto, o de São Tomé e Príncipe, com índices de escolarização em português próximos dos 100%, e onde a emergência de uma variedade nacional é aceite e reconhecida como fator positivo e determinante para a consolidação da identidade nacional.

Existem muitas outras línguas pluricêntricas espalhadas pelo mundo (e.g. espanhol, francês, inglês, árabe, suaíli, chinês). A forma como as autoridades dos vários países envolvidos gerem a pluricentricidade (entendida como “qualidade objetiva de (ser/o que é) pluricêntrico”) é que é diversa. No entanto, a existência de acordos multilaterais de gestão linguística, particularmente ortográficos, não é um exclusivo da língua portuguesa; bem pelo contrário, diversas comunidades ligadas por línguas pluricêntricas têm, nas últimas décadas, desenvolvido este tipo de cooperação multilateral, até como forma de obstar à sempre possível desagregação de uma língua em várias diferentes.

O arranque dos projetos pluricêntricos do IILP coincidiu com renovado interesse pelo estudo das línguas pluricêntricas a nível internacional, que se traduziu e.g. na criação do International Working group on Non-dominant Varieties of Pluricentric Languages¹⁰, em 2010, e na organização de diversas conferências ao longo dos anos. A de 2013 teve a língua portuguesa como eixo temático e dela participaram José Pedro Ferreira e a autora, com a apresentação de uma comunicação extensa, infelizmente inédita. Não surpreende, portanto, a associação dos conceitos de ‘pluricentricidade’ e de ‘pluricentrismo’ (entendido como “sistema de conceitos / ação política que visa a promoção do caráter pluricêntrico ou da pluricentricidade”) ao modelo de gestão partilhada e multilateral da língua portuguesa e aos projetos em apreço neste artigo¹¹.

Os conceitos de pluricentricidade e pluricentrismo foram perfilhados pelo IILP e têm feito, desde então, curso no discurso político da CPLP e de países como Brasil, Moçambique e Portugal.

⁹ A este respeito, leia-se Miguel (no prelo).

¹⁰ Cf. <https://pluricentriclanguages.org/pluricentricity/what-is-a-pluricentric-language/>

¹¹ Sem chamar ao português ‘língua pluricêntrica’ e ao modelo de gestão previsto ‘pluricentrismo’, os Estatutos do Instituto Internacional de Língua Portuguesa, aprovados em 2010, trazem implícitos estes conceitos.

3. Factos determinantes para a afirmação do IILP

3.1. O Acordo Ortográfico da Língua Portuguesa (1990)

O Acordo Ortográfico da Língua Portuguesa, de 1990 (de agora em diante, AOLP90) foi subscrito pelos então sete países de língua oficial portuguesa (Angola, Brasil, Cabo Verde, Guiné-Bissau, Moçambique, Portugal e São Tomé e Príncipe), a que se juntou, já em 2004, Timor-Leste, após a conquista e consolidação da sua independência¹². O seu principal objetivo foi sanar o cisma ortográfico gerado por Portugal em 1911, ao aplicar uma reforma ortográfica sem consultar o Brasil, e artificialmente alimentado pelas academias de ambos os países (com realce para a Academia das Ciências de Lisboa) por mais de um século (cf. Correia et al., 2020).

O AOLP90 fundamentou-se na crença, até então sustentada pela realidade, da existência de duas variedades principais do português, a “brasileira” e a “lusó-afro-asiática” (mais tarde também oceânica)¹³, visão ainda hoje remanescente na mente daqueles que não reconhecem o inevitável desenvolvimento de normas nacionais do português em África e/ou rejeitam a codificação dessas variedades pelas autoridades dos respetivos países.

No que respeita à realização do instrumento ‘vocabulário ortográfico comum’, estipula o artigo 2.º do Tratado do AOLP90:

Artigo 2.º

Os Estados signatários tomarão, através das instituições e órgãos competentes, as providências necessárias com vista à elaboração, até 1 de Janeiro de 1993 [sic], de um vocabulário ortográfico comum da língua portuguesa, no que se refere às terminologias científicas e técnicas.

Desta formulação fica claro que, por um lado, para os seus autores não existia uma clara distinção entre a natureza e as funções de um ‘vocabulário ortográfico’ e de uma ‘terminologia científica e técnica’, e, por outro, não existia definição do modelo do que seria um ‘vocabulário ortográfico comum’. A redação infeliz do Artigo 2.º contribuiu para instalar confusão entre as autoridades dos Estados-membros da CPLP, o IILP e as suas Comissões Nacionais representadas no Conselho Científico do Instituto.

Importa frisar, no entanto, que o texto do Acordo poderá ser considerado um primeiro movimento no sentido de promover o reconhecimento das diferentes variedades do português, ao integrar todos os países de língua oficial portuguesa, mas é, como não poderia deixar de ser, um texto do seu tempo, que hoje, passados mais de 30 anos da sua assinatura, denota envelhecimento. De qualquer forma, foi um pretexto – desperdiçado por alguns – para discutir ortografia, questionar a “pertença” da língua, reconhecer e afirmar a existência de variedades nacionais do português que não a portuguesa e a brasileira, uma

¹² A Guiné Equatorial viria a aderir à Comunidade de Países de Língua Portuguesa em 2014, mas não subscreveu o AOLP90 e não participou dos trabalhos dos projetos do IILP; acreditamos, por o português ser neste país língua oficial apenas de jure, situação que não conheceu até hoje qualquer desenvolvimento.

¹³ A este propósito, veja-se a entrevista concedida por João Malaca Casteleiro ao jornal El País Brasil, em 8 de maio de 2015 (https://brasil.elpais.com/brasil/2015/05/07/cultura/1431024107_637253.html), bem como a entrevista concedida por Evanildo Bechara ao Jornal de Letras (Portugal), na edição de 20 de fevereiro a 5 de março de 2013.

excelente oportunidade para definir as relações entre as essas variedades nacionais e, acima de tudo, para produzir recursos linguísticos computacionais e/ou disponíveis gratuitamente online, para todas elas, de um modo responsável, democrático e partilhado, com respeito efetivo por todos os países da CPLP.

3.2. O Plano de Ação de Brasília, de 2010

O Plano de Ação de Brasília, de 2010 (de agora em diante, PAB2010) constitui um documento anexo à resolução da VI Reunião dos Chefes de Estado e de Governo da CPLP, realizada em Brasília, em 31 de março de 2010¹⁴. Este plano estabelece estratégias várias e linhas de ação para a internacionalização da língua portuguesa e constitui, na nossa opinião, um documento fundador daquilo que viria a ser a nova visão de gestão da língua portuguesa, partilhada entre todos os Estados-membros da CPLP, sob coordenação do IILP. Outros autores, tais como Oliveira (2016), Mendes (2019) e Pinto (2021), partilham desta opinião. A terceira linha de ação do documento, intitulada “III - Estado de desenvolvimento do Acordo Ortográfico (...)” estipula, no seu ponto 3.,

O acordo ortográfico da língua portuguesa não prevê a elaboração de vocabulário ortográfico comum, mas apenas de um vocabulário comum das terminologias científicas e técnicas, cuja existência, nos termos do acordo, não está posta como condição prévia à sua aplicação,

Em consequência, no ponto seguinte, determina que

4. (...) As instituições e os organismos responsáveis pela normalização terminológica nos países da CPLP deverão ser congregados em um projeto, para a elaboração, edição e distribuição de glossários, com áreas ou domínios cuja eleição deverá obedecer a critérios que atendam as realidades terminológicas dos estados membros [sic].¹⁵

No entanto, no ponto 6. da mesma linha de ação, o PAB determina que:

É recomendável que os estados membros que ainda não possuem vocabulário ortográfico da língua portuguesa desenvolvam a sua elaboração, edição e distribuição, de modo que se produza, posteriormente, em formato eletrónico, sob coordenação do IILP e com a participação de uma comissão de especialistas designados pelos estados membros, o vocabulário ortográfico comum da língua portuguesa, que consolide tanto o léxico comum quanto as especificidades de cada país.

Este ponto 6. é crucial para avaliar as consequências benéficas que as decisões do PAB2010 tiveram para os posteriores desenvolvimentos. Este ponto não só contraria o que havia sido afirmado no ponto 4., como:

¹⁴ Ao Plano de Ação de Brasília seguiram-se os de Lisboa (2013), de Díli (2018) e da Praia (2021), mas nenhum destes teve o mesmo impacto que o primeiro.

¹⁵ Esta observação é tanto mais insólita e inexequível quanto não existe em nenhum dos países da CPLP, Brasil e Portugal incluídos, nenhum organismo responsável pela normalização terminológica, nem público, nem privado. Esta falta constitui, de resto, um dos pontos em que a ausência de uma política linguística clara e consubstanciada se faz sentir de forma mais gritante e lesiva para a língua portuguesa.

- por um lado, lança um primeiro esboço da forma de produzir um vocabulário ortográfico comum, determinando que a cada país que ainda não possui um vocabulário ortográfico nacional (VON) compete a sua “elaboração, edição e distribuição”,
- por outro lado, atribui ao IILP, com a participação de uma comissão de especialistas designados pelos estados-membros, a prerrogativa de produzir “o vocabulário ortográfico comum da língua portuguesa, que consolide tanto o léxico comum quanto as especificidades de cada país”;
- determina o caráter integralista (incluindo, para cada país, o vocabulário comum a todos e o que é específico), e não contrastivo, dos VON.

4. A génese dos projetos pluricêntricos do IILP

Neste ponto, pretende-se descrever a génese e desenvolvimento daquele que consideramos o primeiro grande projeto pluricêntrico do IILP, i.e., o VOC, que desempenhou um papel seminal na definição da que seria a política linguística desenvolvida pelo IILP ao longo dos últimos 13 anos. Referência será feita aos projetos PPPLE e TCTC, por estes terem profundas relações com o VOC.

4.1. O Vocabulário Ortográfico Comum da Língua Portuguesa (VOC)

O VOC constitui o primeiro vocabulário ortográfico comum da língua portuguesa. A sua primeira edição data de maio de 2017 e inclui os vocabulários ortográficos nacionais de cinco Estados-membros da CPLP: Brasil, Cabo Verde, Moçambique, Portugal e Timor-Leste¹⁶. A relevância histórica deste instrumento reside, entre outros aspetos, no facto de ter sido o primeiro vocabulário ortográfico para o português a alinhar, com base em interpretação comum do AOLP90, os vocabulários de Brasil e Portugal – v. Correia et al., 2020 – e, além disso, ter constituído a primeira recolha sistemática do vocabulário em uso no português falado em Cabo Verde, Moçambique e Timor-Leste.

Apesar de não ter conhecido publicação impressa, prevê-se a publicação digital da sua 2.^a edição até ao final de 2022, em diferentes plataformas abertas e com formatos de exportação atualizados.

4.1.1. Os primeiros passos

Em maio de 2009, o Instituto de Linguística Teórica e Computacional (ILTEC)¹⁷, submeteu ao então Fundo Especial da Língua Portuguesa¹⁸ um projeto para a elaboração do Vocabulário Ortográfico do Português (VOP), que previa duas fases:

¹⁶ Embora se encontre recolhida uma primeira versão desde 2014 do VON de São Tomé e Príncipe, aguarda-se a sua conclusão e validação pelas autoridades nacionais, para que possa ser integrado e disponibilizado na plataforma VOC. A sua cedência oficial ao IILP data de 2019. Na reunião ordinária do Conselho Científico do IILP de 2022, foi anunciada a elaboração de um VON da Guiné-Bissau, cuja equipa nacional terá, para o efeito, apoio do Ministério das Relações Exteriores do Brasil.

¹⁷ Unidade de investigação e desenvolvimento com financiamento da Fundação para a Ciência e a Tecnologia, fundada em 1987 e extinta em 2015, em virtude da sua fusão com o Centro de Linguística Geral e Aplicada da Universidade de Coimbra, dando origem ao CELGA-ILTEC.

¹⁸ O Fundo Especial da Língua Portuguesa foi criado durante a vigência do XVII Governo Constitucional da República Portuguesa e foi gerido pelo IPAD (Instituto Português para a Cooperação e o Desenvolvimento), sobre o qual se encontra informação aqui: <https://www.instituto-camoes.pt/sobre/sobre-nos/gestao-sobrenos/historia/historia-do-ipad>.

- execução de um vocabulário ortográfico eletrónico para Portugal;
- coordenação de um vocabulário ortográfico comum, envolvendo equipas de todos os países da CPLP.

O VOP foi, então, concebido como: a) um vocabulário representativo do português europeu, mas com indicação da variação ortográfica entre Brasil e Portugal, e b) um serviço público, de acesso universal e gratuito, mas c) uma plataforma digital, atualizável e extensível, podendo ser usada por outras equipas, por exemplo, para a elaboração do que veio a ser o VOC – cf. Correia et al. (2014).

Esta conceção fundava-se, por um lado, na necessidade de elaborar um “vocabulário ortográfico comum” e, por outro, na constante preocupação do ILTEC e da autora em produzir ferramentas computacionais interoperáveis e reutilizáveis (cf. e.g. Correia, 1994), bom como de fomentar sinergias entre os países de língua portuguesa, de modo a produzir recursos linguísticos em conjunto (cf. Correia, 1998).

Em junho de 2010 foi publicada a primeira edição do VOP (cf. Correia et al., 2014), em formato eletrónico, produzido pelo ILTEC, recurso que viria a tornar-se, a par do Lince – Conversor para a Nova Ortografia (Ferreira et al., 2012b), instrumento oficial para aplicação do Acordo Ortográfico em Portugal, pela sua aprovação por unanimidade pela Assembleia da República Portuguesa (Deliberação n.º 3-PL/2010, de 15 de dezembro), bem como pela publicação da Resolução do Conselho de Ministros 08/2011, de janeiro de 2011.

Por seu turno, o Brasil, através da Academia Brasileira de Letras (ABL), havia publicado em 2008, em volume impresso, a 5.^a edição do seu Vocabulário Ortográfico da Língua Portuguesa (VOLP), incluindo as alterações resultantes da sua interpretação do AOLP90. Os demais países signatários do AOLP90 não possuíam ainda os respetivos vocabulários ortográficos.

Recém-empossado como Diretor Executivo do IILP, Gilvan Müller de Oliveira tomou conhecimento da existência do VOP e, em outubro de 2010, contactou a Direção do ILTEC e os coordenadores do Vocabulário, no sentido de indagar da disponibilidade do Instituto e dos seus especialistas computacionais para colaborar com o IILP no cumprimento das incumbências que lhe haviam sido cometidas pelo PAB2010, i.e., a execução do vocabulário ortográfico comum previsto no texto do AOLP90.

Em dezembro de 2010, o ILTEC foi convidado a participar na Reunião Extraordinária do Conselho Científico do IILP, que teve lugar em Brasília, de modo a apresentar ao Conselho o modelo de vocabulário ortográfico comum, concebido por José Pedro Ferreira e pela autora. Tal proposta teve por base o raciocínio já desenvolvido e registado aquando da submissão do projeto VOP ao Fundo Especial da Língua Portuguesa, em 2009.

4.1.2. A conceção do projeto

A apresentação feita durante a Reunião Extraordinária do Conselho Científico do IILP teve como suporte uma apresentação em PowerPoint e foi consubstanciada em um documento de 10 páginas – Correia e Ferreira (2011)¹⁹ –, entregue ao então Diretor Executivo do IILP, a seu pedido, em abril de 2011.

¹⁹ Ambos os documentos fazem parte dos arquivos do IILP.

O primeiro objetivo dos autores, na apresentação em 2010 e no documento de 2011, foi dirimir a confusão que se havia estabelecido a propósito das redações infelizes quer do AOLP90, quer do PAB2010. Com esse objetivo, começou-se por delimitar os conceitos em apreço, definindo, em primeiro lugar, ‘vocabulário ortográfico comum’ do seguinte modo:

Atendendo aos exemplos de que dispomos para a língua portuguesa, um vocabulário ortográfico é uma lista de unidades lematizadas (representadas pela sua forma de citação) que esclarece a sua forma ortográfica, a sua categoria e contém, ainda, para algumas delas, indicações de idiosincrasias que se prendem com a sua pronúncia e/ou flexão. No caso do VOP, graças à sua estrutura computacional, apresentam-se ainda todas as formas flexionadas associadas a cada lema, bem como a sua divisão silábica, marcação da sílaba tónica e de relações morfológicas entre entradas diferentes. O objetivo principal de um vocabulário ortográfico é o de veicular uma norma relativamente às palavras de uma língua: à sua ortografia, pronúncia e flexão. Um vocabulário ortográfico comum para o âmbito dos países da CPLP é um vocabulário que integra não apenas o vocabulário que é comum a todas as variedades nacionais da língua portuguesa, como ainda os vocabulários específicos de cada uma dessas variedades. Deve, portanto, conter todas as palavras que são comuns (e que não serão portadoras de qualquer marca de uso) e ainda todas as palavras que são específicas de cada variedade (essas sim, portadoras de marca geográfica).

Correia e Ferreira, 2011

Em seguida, definiram-se terminologias científicas e técnicas, do seguinte modo:

‘Terminologias científicas e técnicas’ é uma expressão que tem dois significados à partida:

- i. termos de diferentes domínios de especialidade, que caracterizam o discurso de cada um desses domínios; trata-se em suma de palavras da língua que, em discursos especializados, assumem comportamento específico, em particular significado e denominação específicos;
- ii. produtos lexicográficos especializados, também conhecidos como dicionários terminológicos, que registam e descrevem os termos das diferentes especialidades.

Correia e Ferreira, 2011

Seguindo a mesma linha de pensamento, procurou-se deixar explícitos os condicionalismos impostos à realização de ‘terminologias comuns’, esclarecendo que:

Tal como as palavras da língua corrente, as terminologias científicas e técnicas, na acepção i., também variam entre variedades linguísticas. O número de terminologias específicas dentro de cada língua é praticamente ilimitado e depende dos critérios usados para a sua delimitação. As terminologias científicas e técnicas carecem de ser fixadas e descritas através de glossários e dicionários, até porque, infelizmente, nenhuma variedade do português está dotada de descrições terminológicas suficientes.

Correia e Ferreira, 2011

Posto isto, foi apresentado um conjunto de necessidades a colmatar para se poder avançar com a execução de terminologias científicas e técnicas no espaço da CPLP. O IILP precisava então (e precisa até hoje) de:

- delimitar os domínios a descrever;
- definir o peso relativo de cada domínio no âmbito de uma ferramenta comum;
- delimitar as áreas prioritárias de intervenção;
- nomear comissões nacionais de especialistas de domínio e de linguistas;
- definir as características da ferramenta terminológica comum;
- construir / desenvolver a plataforma digital de trabalho comum²⁰.

Correia e Ferreira, 2011

Após a apresentação ao Conselho Científico (dezembro de 2010) e da entrega do projeto escrito em 2011, foi deixada ao Conselho Científico do IILP e à sua Direção Executiva (DE) a prerrogativa de decidir se lhe importava:

- avançar com a produção do VOC,
- avançar com o desenvolvimento de terminologias científicas e técnicas,
- avançar com ambos os projetos em simultâneo.

O ILTEC declarou possuir condições de colaborar no desenvolvimento de ambos os projetos, mas manifestou ao Conselho a sua convicção de que, naquela época, seria mais importante avançar com a execução do VOC (exequível no imediato) e amadurecer o que se pretendia exatamente com as terminologias científicas e técnicas do espaço CPLP (Correia e Ferreira, 2011).

É conhecida a decisão tomada pelo Conselho Científico pela DE de avançar, então, com a execução do VOC, reunindo condições para, posteriormente, avançar com as TCTC.

4.1.3. Os primeiros passos do VOC

Em 2011, decorreu um curso sobre políticas linguísticas para o português e sobre o VOC, na sequência da Reunião Ordinária do Conselho Científico do IILP, destinada às potenciais equipas nacionais de especialistas (EN), na qual foram apresentadas três propostas para a execução do VON de cada país, que já haviam, de resto, sido adiantadas em Correia e Ferreira (2011), documento que reza:

O VOC será construído com base na estrutura e nas ferramentas do Portal da Língua Portuguesa (PLP)²¹, que alojará o recurso e a ele permitirá acesso de forma gratuita e livre. cada país pode contribuir de diferentes formas:

²⁰ A incapacidade do IILP de tomar as ações referidas, e outras, encontra justificação, entre outros, nos seguintes fatores: i) indefinição jurídica, e.g. relativamente aos estatutos do Instituto, de que existem duas versões assaz contraditórias aprovadas pelo Conselho Científico (em 2005 e em 2010); ii) subfinanciamento crónico do Instituto, especialmente devido ao não pagamento das contribuições obrigatórias por um conjunto significativo de países; iii) falta de participação sistemática nos trabalhos do IILP, por parte de alguns países; iv) pouca adequação da constituição das CN, que em conjunto constituem o Conselho Científico; v) desinteresse dos governos nacionais, que não dotam as suas CN de meios e recursos que lhes permitam desenvolver as atividades acordadas nas reuniões do Conselho Científico do IILP; vi) mudanças bruscas de rumo e modo de atuação, provocadas pela substituição governos e de altos funcionários; vii) incapacidade ou falta de vontade dos Estados-membros da CPLP de compreender o potencial de desenvolvimento do IILP e a sua importância para o futuro da Comunidade.

²¹ <http://www.portaldalinguaportuguesa.org/vop.html>

- a) nomeando um vocabulário existente como representante da sua variedade nacional;
 - b) nomeando uma equipa nacional que, usando as ferramentas lexicográficas do PLP, construa um VON [vocabulário ortográfico nacional];
 - c) nomeando as fontes que devem ser usadas pela equipa central do PLP para criar um VO [vocabulário ortográfico] e nomeando uma comissão de acompanhamento, que, no seio do IILP, confira legitimidade ao recurso final, avaliando a adequação da descrição conseguida à sua variedade nacional.
- Qualquer que seja a opção tomada, os conteúdos de cada VON serão integrados numa plataforma comum, constituindo-se um VOC (...)"

No mesmo momento (agosto de 2011), o ILTEC cedeu formalmente ao IILP o seu VOP²² e a plataforma computacional respetiva, de modo a servir de base à construção do VOC. Também o Brasil, na pessoa de Evanildo Bechara, cedeu formalmente ao IILP a 5.^a edição do Vocabulário Ortográfico da Língua Portuguesa da Academia Brasileira de Letras, para constituir o VON do Brasil. Os demais países elaborariam de raiz os seus vocabulários, de acordo com princípios e especificações comuns, com base em dados de uso extraídos de corpora, para também serem integrados na plataforma comum. Dadas as dificuldades de alguns destes países para o desenvolvimento dos seus recursos nacionais, as equipas nacionais optaram por realizar o trabalho com apoio técnico direto da equipa central (EC).

Ainda em 2011, decorreu na Praia a Primeira Formação Técnica das equipas nacionais do VOC, dirigida por José Pedro Ferreira, da Equipa Central.

4.1.4. A estrutura de funcionamento

Como projeto internacional e comum a um conjunto de países, o VOC foi um projeto coordenado pelo IILP e organizado em torno de três corpos: a Equipa Central (EC), as Equipas Nacionais (EN) e o Corpo de Consultores. Mutatis mutandi, este modelo veio a ser assumido, a partir de então, pela generalidade dos projetos do IILP e nunca suscitou dúvidas às Comissões Nacionais do IILP²³.

A EC²⁴ teve por missão pôr em execução o projeto, formar as equipas nacionais, produzir a documentação necessária²⁵, adaptar a plataforma digital para acolher os VON, coordenar e apoiar o trabalho das equipas nacionais na constituição dos corpora e dos VON, centralizar e harmonizar os dados dos vários países, preparar a publicação e levar a cabo a gestão da plataforma. Esta equipa foi constituída ad hoc pelos especialistas em lexicografia

²² Ainda que o VOP tivesse sido financiado por fundos públicos, a propriedade do recurso foi do ILTEC e posteriormente do CELGA-ILTEC, não tendo até hoje o Estado Português concretizado a sua posse, como lhe tem sido e proposto repetidamente pelos seus executores.

²³ É, portanto, com perplexidade e desagrado que se consideram todas as questões levantadas relativamente ao modelo de gestão de projetos do IILP, 11 anos depois de inaugurado, durante o Curso de Formação para o desenvolvimento do projeto das TCTC no espaço da CPLP (Brasília, outubro de 2021).

²⁴ Constituída inicialmente por José Pedro Ferreira, Margarita Correia e Gladis M. B. Almeida, veio a integrar, posteriormente, Inês Machungo, coordenadora da EN de Moçambique, devido ao seu vasto conhecimento de lexicografia e à forma modelar como esta EN se envolveu e contribuiu de forma determinantes para o sucesso do projeto. Margarita Correia abandonou a EC em maio de 2018, após ter sido eleita Presidente do Conselho Científico do IILP, tendo retomado a sua função no projeto em junho de 2022.

²⁵ E.g. relatórios de progresso e final, memorandos a apresentações públicas, artigos e comunicações em eventos científicos e, ainda, Equipa Central do VOC / IILP. 2012.

e em lexicografia computacional que idearam e executaram as funções a ela atribuída. Nem a forma como esta equipa foi gerada, nem o modelo, nem a sua constituição e nem as suas atribuições foram até hoje alvo de reparos em reuniões do Conselho Científico do IILP.²⁶

Às EN, nomeadas pelas autoridades dos respetivos países e constituídas por técnicos especializados a trabalharem localmente, competiu recolher, selecionar e tratar as fontes usadas para a construção dos VON, compilar e rever as listas de palavras extraídas dos corpora ou das fontes (impressas ou outras) de produção nacional escritas em português, finalizar e validar junto das autoridades dos seus países a versão final a ser entregue ao IILP. Um representante da EC realizou uma missão a cada país, para apoio técnico.

Às EN foi cometido trabalho técnico de lexicografia; à EC, além do trabalho de lexicografia, também o de programação e gestão computacional²⁷.

Dada a relevância e o impacto político do VOC, que visou aplicar a norma ortográfica emanada do texto do AOLP90, foi constituído um Corpo Internacional de Consultores, constituído por especialistas em ortografia, lexicografia e/ou política linguística de cada Estado-membro e por ele nomeados²⁸, para discussão de questões técnicas. A sua tarefa consistiu em estabelecer entendimento comum quanto à aplicação do AOLP90, identificando e resolvendo dissensões e lacunas de modo a chegar a uma interpretação consensuada do texto do tratado (cf. Equipa Central do VOC/IILP 2016).

4.1.5. O financiamento do VOC

Ao longo de todo o processo de execução do VOC, e também das TCTC, o ILTEC e posteriormente o CELGA-ILTEC, cedeu graciosamente ao IILP: a) o saber, a experiência e o trabalho dos seus peritos em política linguística e em lexicografia computacional; b) o modelo e a estrutura de ambos os projetos; c) os dados do VOP; d) a estrutura computacional OSLIN, usada para a realização do VOP e as adaptações necessárias para a tornar a base do VOC; e) a formação das Equipas Nacionais (EN) dos projetos e apoio na recolha e inserção dos dados na plataforma VOC; f) grande parte da coordenação técnica do projeto, assegurada por José Pedro Ferreira e Margarita Correia, mas também por Gladis Almeida (UFSCar) e Inês Machungo (UEM); g) apoio logístico à Direção executiva do IILP em todos os eventos e momentos de prestação de contas que ocorreram entre 2010 e 2022.

A Equipa Central foi apoiada por bolseiros de iniciação científica do GETerm, sediado na Universidade Federal de São Carlos, UFSCar, Brasil, coordenados por Gladis Maria de Barcellos Almeida. Além disso, entre maio de 2012 e junho de 2015, foram atribuídas ao ILTEC duas bolsas mensais do Programa Fernão Mendes Pinto, que comparticiparam no pagamento da remuneração de dois bolseiros dedicados ao projeto VOC²⁹.

O IILP recebeu, ainda, 40 mil euros, doados pela Comissão Nacional de Angola, em 2012. Cada Estado-membro participante no projeto garantiu o financiamento das suas respetivas EN e dos seus representantes no Corpo de Consultores.

²⁶ Extemporânea e inexplicavelmente, foi questionado por uma Comissão Nacional, em novembro de 2021, o modelo em vigor há mais de dez anos no IILP, a propósito da constituição da equipa central do projeto TCTC, projeto que tivera início em 2015, quando recebeu a primeira tranche do financiamento obtido junto do Fundo Especial da CPLP.

²⁸ Cf. a ficha técnica do VOC neste endereço: <<http://voc.cplp.org/?action=ficha>>.

²⁹ O valor cada uma destas bolsas era de 500,00 Euros / mês, equivalente ao salário mínimo nacional.

Nenhuma das instituições portuguesas e brasileiras que participaram na execução do projeto nunca recebeu qualquer compensação pecuniária do IILP ou da CPLP, bem como nenhum membro da EC do VOC foi remunerado pelo seu trabalho no projeto.

4.2. As relações do VOC com o PPLE e as TCTC

O projeto Portal do Professor de Português Língua Estrangeira / Língua Não Materna (PPPLE) foi o primeiro projeto-irmão do VOC e teve início em 2011. Tal como aconteceu com o VOC, o PAB2010 cometeu ao IILP “a tarefa de criar uma Plataforma comum na internet, relativa ao ensino do português”. Para esse fim, foi estabelecido um acordo de cooperação técnica com a Sociedade Internacional de Português Língua Estrangeira (SIPLE), para a identificação de especialistas dos diversos países que integram a comunidade da CPLP, para o planeamento e o desenvolvimento do Portal, juntamente com as Comissões Nacionais e o corpo técnico do IILP. De acordo com a informação disponibilizada no Portal (cf. <https://ppple.org/o-portal>), este é “uma plataforma on-line, que tem como objetivo central oferecer à comunidade de professores e interessados em geral, recursos e materiais para o ensino e a aprendizagem do português como língua estrangeira / língua não materna.”

Assim como o VOC, o “Portal é concebido, desenvolvido, alimentado e gerido de forma multilateral, funcionando como instrumento de cooperação linguístico-cultural entre os Estados-Membros da Comunidade de Países de Língua Portuguesa (CPLP). Desse modo, o PPPLE tem um papel muito importante para as estratégias de promoção, difusão e projeção do português no mundo, criando um sistema internacionalizado de gestão do ensino de PLE/PLNM.”

O PPPLE constituiu, tanto quanto sabemos, o primeiro projeto sistematicamente desenvolvido para colmatar a falta de materiais para o ensino de diferentes variedades da língua portuguesa, além das variedades brasileira e portuguesa. Apesar de não ter merecido, ao longo dos anos, o reconhecimento, a atenção e o apoio merecidos por parte de todos os Estados-membros, o PPPLE constitui um dos projetos pluricêntricos do IILP com maior impacto a nível internacional.³⁰

Em 2015, durante o mandato de Marisa Mendonça como Diretora Executiva do IILP, surgiu a possibilidade de propor as TCTC para financiamento ao Fundo Especial da CPLP. Foi a primeira vez na sua história que o IILP recebeu financiamento externo específico para o desenvolvimento de um projeto. Em virtude da cadência de financiamento adotado pelo Fundo Especial, o projeto conheceu duas etapas: a primeira, entre 2015 e 2017, e a segunda, de 2020 ao final de 2022. As duas versões que o projeto conheceu até conseguir o seu primeiro financiamento foram escritas por Gladis M. de Barcellos Almeida, José Pedro Ferreira e Margarita Correia, investigadores que constituíram, quando este arrancou e até 2017, a respetiva Equipa Central. Na sua segunda etapa, o projeto foi coordenado, durante o ano de 2020, por Tanara Zingano Kuhn, bolseira do programa de Bolsas de Cientista Convidado do IILP, tendo Gladis Almeida reassumido a coordenação do mesmo, no início de 2021.

³⁰ Sobre o PPPLE, além do Portal propriamente dito, cf. Mendes (2016) e Furtoso et al. (2018), bem como o número 4 (2018) da revista *Platô*, inteiramente dedicado a este projeto (<https://www.yumpu.com/pt/document/view/62886263/plato-vol-4-n-7-2018-o-portal-do-professor-de-portugues-lingua-estrangeira-lingua-nao-materna-ppple-analises-e-projecoes>). Cf. também Jesus et al. (2018), Mendes (2019) e a entrevista concedida pela autora, Edleise Mendes, à revista da Associação de Professores de Português (Correia e Aido, orgs., 2021).

O projeto TCTC constituiu a primeira tentativa sistemática alguma vez levada a cabo de recolher e harmonizar terminologias científicas e técnicas em uso em cada um dos Estados-membros da CPLP. Infelizmente, porém, a Comunidade no seu todo não possui ainda maturidade científica nem política para reconhecer a importância crucial deste (e de outros) projeto para o cumprimento da missão do IILP e para o desenvolvimento e difusão da língua³¹.

5. Considerações finais

Com este texto, pretendi contribuir para narrar a história do pluricentrismo da língua portuguesa e a do IILP, entidade a quem foi cometida a árdua tarefa de proceder de pôr em prática uma gestão partilhada e democrática da língua portuguesa. A escolha do tema teve motivações várias, das quais destaco a transparência e honestidade intelectual com que sempre procurei pautar a minha atividade profissional, o empenho em que a língua portuguesa seja gerida de modo igualitário e democrático por todos os países que a assumiram como língua oficial e fator de construção da sua identidade nacional, a necessidade de esclarecer factos, desfazer mitos e meias-verdades e contribuir para o diálogo sobre e em língua portuguesa. Tive a preocupação de ser justa e dar o seu a seu dono; espero tê-lo conseguido.

Todas as afirmações e opiniões emitidas neste texto são da minha inteira e exclusiva responsabilidade.

Com a publicação deste e do outro texto de minha autoria neste número da Platô, assim como com a conclusão, até ao final deste ano, dos projetos em que ainda estou envolvida (VOC e TCTC), encerro a minha participação ativa no IILP, cedendo o lugar a quem queira agora assumir responsabilidades no Instituto, a bem desta língua que nos une muito mais do que nos separa e das sociedades que a falam.

6. Agradecimentos

Não posso deixar de aproveitar esta oportunidade para agradecer publicamente aos meus parceiros nesta aventura, José Pedro Ferreira, Gladis Almeida, Inês Machungo e Tanara Zingano Kuhn, pelo tanto que crescemos e aprendemos juntos, pela partilha constante de êxitos e alegrias e pelo apoio nos momentos de fracasso e tristeza.

Obrigada, ainda, à Denise Fonseca, à D. Maria José Sousa e à Nadine Rodrigues, a minha família cabo-verdiana.

Bem hajam todos.

³¹ A oposição a este projeto, assumida desde sempre pela CN de Angola e, desde novembro de 2021, também pela de Portugal, ancorada em supostas razões teórico-metodológicas, denota desconhecimento do que é o trabalho em lexicografia e em terminologia, assim como uma falta de respeito relativamente ao trabalho desenvolvido pelo IILP e seus parceiros.

7. Bibliografia

Almeida, G. M. B. / Ferreira, J. P. / Correia, M. / Oliveira, G. M. (2013, jan-abr): «Vocabulário Ortográfico Comum (VOC): constituição de uma base lexical para a língua portuguesa», *Estudos Linguísticos*, 42(1): p. 204-215. Em linha: <<https://revistas.gel.org.br/estudos-linguisticos/article/view/1100>>

Clyne, Michael G. ed. 1992. *Pluricentric Languages: Differing Norms in Different Nations*. Berlin, New York: Walter de Gruyter.

Correia, M. 1994. *Bases digitais lexicais na União Europeia*. ms. Disponível em https://www.academia.edu/34325622/Bases_digitais_lexicais_na_Uni%C3%A3o_Europeia.

Correia, M. 1998. *Cooperação entre Portugal e Brasil no domínio da terminologia: a expressão de um desejo*. In: *Terminómetro*, número especial n.º 3: Terminologia no Brasil. Paris: União Latina, pp. 42-43.

Correia, M. 2019. *VOC – conceção, estrutura, realização, resultados*. In: *Dialogues Ibériques - Assises de la Langue Espagnole et de la Langue Portugaise*. Faro: Imprensa da Universidade do Algarve.

Correia, M.; Ferreira, J. P. 2011. *Vocabulário Ortográfico Comum – VOC, Proposta de execução do ILTEC*. (ms.) Lisboa: ILTEC, 21 de abril de 2011.

Correia, M.; Ferreira, J.P. 2014. *Vocabulário Ortográfico do Português: descrição do projeto*. In: Garcés Gómez, María Pilar (ed.). *Lexicografía teórica y aplicada* (Anexo de Revista de Lexicografía), ed. A Coruña: Universidade da Coruña, pp. 91 -102

Correia; Ferreira; Almeida. 2020. *A gestão da ortografia da língua portuguesa: do desencontro ao Vocabulário Ortográfico Comum da Língua Portuguesa*. *Estudis Romànics*, [en línia], vol. 42, p. 277-86, <https://raco.cat/index.php/estudis/article/view/370606>

Correia, M.; Aido, J.P. orgs. 2021. *Palavras – revista em linha*, n.º 4 – 2021. Disponível em <https://palavras.appform.pt/ojs/index.php/revista/index>

Equipa Central do VOC / IILP. 2012. *Manual para a elaboração de corpora com vista à organização dos Vocabulários Ortográficos Nacionais dos países integrantes da CPLP*. Praia: IILP (ms.).

Equipa Central do VOC / IILP. 2016. *Sistematização das Regras de Escrita do Português para Aplicação do Acordo Ortográfico da Língua Portuguesa, de 1990, no âmbito do Vocabulário Ortográfico Comum da Língua Portuguesa*. Praia: IILP (ms.).

Equipa Central do VOC / IILP (2016): *Sistematização das Regras de Escrita do Português para Aplicação do Acordo Ortográfico da Língua Portuguesa, de 1990, no âmbito do Vocabulário Ortográfico Comum da Língua Portuguesa*. Praia: IILP (ms.).

Faraco, C. A. (2012). *Lusofonia: utopia ou quimera? Língua, história e política*. In: In Lobo, T.; Carneiro, Z.; Soledade, J., Almeida, A.; Ribeiro, S., orgs. *Rosae: linguística histórica, história das línguas e outras histórias* [online]. Salvador: EDUFBA

Faraco, C. A. (2016): *História Sociopolítica da Língua Portuguesa*. S. Paulo: Parábola.

Ferreira; Almeida; Correia; Oliveira. 2012a. *The Common Orthographic Vocabulary of the Portuguese Language: A Set of Open Lexical Resources for a Pluricentric language*. In: *Proceedings of the Eighth International Conference on Language Resources and evaluation (LREC'12)*. Istanbul: European Language Resources Association (ELRA) pp. 1071-1075.

Ferreira, J.P.; Lourinho, A.; Correia, M. 2012b. *Lince, an End User Tool for the Implementation of the Spelling Reform of Portuguese*. In: Helena Caseli; Aline Villavicencio; António Teixeira; Fernando Perdigão (eds.). *Computational Processing of the Portuguese Language*. ISBN: 978-3-642-28884-5. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 46-55. doi: 10.1007/978-3-642-28885-2_5

Ferreira, J. P.; Janssen, M.; Almeida, G. B.; Correia, M.; Oliveira, G. M. (2012): «*The Common Orthographic Vocabulary of the Portuguese Language: a set of open lexical resources for a pluricentric language*», in Calzolari, N. et al. (eds.): *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*. s/l: ELRA. Em linha: <http://www.lrec-conf.org/proceedings/lrec2012/pdf/1034_Paper.pdf> [Consulta: 22 de agosto de 2019].

Ferreira, J. P.; Correia, M.; Almeida, G. B. (orgs.) (2017): *Vocabulário Ortográfico Comum da Língua Portuguesa*. Praia: Instituto Internacional da Língua Portuguesa / Comunidade dos Países de Língua Portuguesa. Em linha: <<http://voc.cplp.org/index.php?action=apresent>>.

Furtoso, V.V.; Araújo, V.C; Killner, R.M. (2017). *As potencialidades do Portal da Professora de Português Língua Estrangeira / Língua Não Materna como Recurso Educacional Aberto*. In: El Kadri, M.S.; Ortenzi, D.I.G.; Ramos, S.G.M. (eds.) *Tecnologias digitais no ensino de línguas e na Formação de professores: reorganizando sistemas educacionais*. Campinas: Pontes. Gonçalves, P. (2010). *A génese do Português de Moçambique*. Lisboa: INCM.

Janssen, M. (2005). *Open Source Lexical Information Network*. In: Bouillon, P.; Kanzaki, K. (eds.) *Proceedings on the Third International Workshop on Generative Approaches to the Lexicon (May 19-21 2005)*. Genève: École de Traduction et d'Interprétation – Université de Genève, p. 79-106.

Janssen, M.; Ferreira, J. P. (2018). *Technical Implementation of the Vocabulário Ortográfico Comum da Língua Portuguesa*. *Computational Processing of the Portuguese Language*. Cham: Springer International Publishing, p. 191-200.

Jesus, P. C. S. G. de; Oliveira, G. M. (2018). *Ensinando línguas em uma perspectiva pluricêntrica: o Portal do Professor de Português Língua Estrangeira/Língua Não Materna (PPPLE)*. *Domínios de Linguagem*, v. 12 n. 2 (2018): Português como Língua Estrangeira, Português

como Língua de Herança, Português como Língua Adicional. <https://doi.org/10.14393/DL34-v12n2a2018-13>

Machungo, I. coord. 2017. *VOMOLP - Vocabulário Ortográfico Moçambicano da Língua Portuguesa*. Maputo: Cátedra de Português Língua Segunda e Estrangeira. Em linha: <https://voc.cplp.org/index.php?action=von&von=mz>

Mendes, E. (2016). *Pluricentrismo linguístico, ensino e produção de materiais de Português LE no PPPLE*. In: Alvarez, M. L. O.; Gonçalves, L. (orgs.). *O Mundo do Português e o Português no Mundo Afora: especificidades, implicações e ações*. Campinas: Pontes, 2016, p. 293-310.

Mendes, E. (2019). *A promoção do português como língua global no século XXI – um cenário a partir do Brasil*. *Linha D'Água* (online). São Paulo, v. 32, n. 2, p. 37-64.

Miguel, A. (no prelo). *Das políticas às práticas: possibilidades da noção de português língua pluricêntrica em formação de professores*. In: Antunes, A. J. et aliae (ogs.). *Português língua pluricêntrica: das políticas às práticas*. Porto Alegre: EDIPUCRS (e-book a ser publicado em 2022).

Muhr, R. 2016. *The state of the art of research on pluricentric languages: Where we were and where we are now*. In: Muhr, R.; Fonyuy, K.E.; Ibrahim, Z.; Miller, C. (eds.) *Pluricentric Languages and non-dominant Varieties worldwide*. Volume 1: *Pluricentric Languages across continents – Features and usage*. Wien et. al., Peter Lang Verlag. p. 9-32.

Neto, F. T. C. 2013. *História Do Futuro: Diagnóstico e perspectivas de políticas públicas para o ensino/aprendizagem de PLE-PL2 no Brasil do século XXI*. Dissertação de mestrado em Linguística Aplicada. Brasília: Departamento de Línguas Estrangeiras e Tradução-LET, UnB.

Oliveira, G. M. (2013). *Política linguística e internacionalização: a língua portuguesa no mundo globalizado do século XXI*. *Trabalhos de Linguística Aplicada*, Campinas, n (52.2): 409-433, jul./dez. 2013

Oliveira, G. M. (2016). *The system of national standards and the demolinguistic evolution of Portuguese*. In: MUHR, R. (ed.). *Pluricentric Languages and Non-Dominant Varieties Worldwide*. Part II: *The Pluricentricity of Portuguese and Spanish*. New Concepts and Descriptions. Frankfurt am Main, Peter Lang GmbH, , p. 35-48.

Pinto, P. F. (2021). *Pluricentrismo em Portugal*. *Palavras em linha*, n.º 4, p. 69-73. Disponível em <https://palavras.appform.pt/ojs/index.php/revista/article/view/138/131>

Pinto, P. F. (no prelo). *Português, língua pluricêntrica: teoria e realidades*. In: Antunes, A. J. et aliae (ogs.). *Português língua pluricêntrica: das políticas às práticas*. Porto Alegre: EDIPUCRS (e-book a ser publicado em 2022).

Variedades nacionais de línguas pluricêntricas: o caso do português em Moçambique

Inês Machungo e Gregório Firmino

Sobre o autor

Inês Machungo - Universidade Eduardo Mondlane, Professora Jubilada, Doutorada em Linguística, machungoines@aol.com

Gregório Firmino - Universidade Eduardo Mondlane, Professor Associado, Doutorado em Antropologia Linguística, gregoriofirmino@gmail.com

Resumo

Em países africanos de dominação colonial, as línguas ex-coloniais expandiram o seu espaço de uso e papel social como entidades simbólicas e meios de comunicação comunitária. Estas variedades não-nativas de línguas coloniais (VNN) (KACHRU, 1986), tal como qualquer outra língua, têm passado por transformações significativas do ponto de vista estrutural e simbólico podendo afirmar-se que a sua nativização (BOKAMBA, 1982; FÉRAL, 2020; LIPSKI, 1985; 2004; MANNESSY, 1989; MESTHRIE; MOHA, 2008; TUNDE-AWE, 2019) é uma realidade.

As manifestações linguísticas diferenciadas que caracterizam estas VNN resultam, em geral, do contacto entre as línguas indígenas, em sociedades multilingues, e as línguas ex-coloniais, do qual se desenvolve um processo de fertilização mútua, enriquecendo as línguas autóctones e as línguas ex-coloniais.

Este texto debruça-se sobre o português de/em Moçambique, mostrando-se a sua inserção no país e os desafios que se lhe colocam, decorrentes das funções que lhe estão associadas: língua oficial e língua de unidade nacional sendo, por essa razão, o português, uma língua franca, a língua falada em todo o território nacional.

O processo de nativização do português em Moçambique compreende uma dimensão simbólica e linguística; simbólica, caracterizada pela emergência de novas atitudes e ideologias sociais face ao uso da língua, e linguística relacionada com as novas formas linguísticas incorporadas na língua (FIRMINO, 2002). Neste texto são exemplificadas algumas dessas mudanças linguísticas, com destaque para o domínio lexical por ser o que se destaca nos processos de nativização (SCHNEIDER 2007: 56) e por ser também aquele que tem sido

objecto de muitos estudos sobre o português de Moçambique (DIAS, 2002; LOPES et al., 2016; MACHUNGO, 2000; 2015; MENDES, 2010; TIMBANE, 2013; 2016 entre outros).

Palavras-chaves

Léxico; línguas pluricêntricas; nativização; português de Moçambique.

1. 1. A problemática das línguas ex-coloniais em países pós-coloniais

A visão sobre a situação das línguas ex-coloniais nos países africanos pós-coloniais é muitas vezes dominada por considerações ideológicas que, embora possam ser justificadas pela postura anti-colonial, falham em reconhecer o seu estado actual. Passadas várias décadas desde o advento das independências em África, as línguas ex-coloniais expandiram o seu espaço e papel sociais, quer como entidades simbólicas quer como meros meios de comunicação. Sob o ponto de vista estrutural têm passado por transformações significativas, aventando-se já a certeza da sua nativização (BOKAMBA, 1982; FÉRAL, 2020; LIPSKI 1985, 2004; MANNESSY, 1989; MESTHRIE; MOHA, 2008; TUNDE-AWE, 2019). Simultaneamente, as línguas autóctones não permaneceram estáticas, ou seja, elas também adquiriram novas funções sociais, em consequência da sua adaptação à dinâmica funcional dos países pós-coloniais.

Como Tengan (1994) referiu, as sociedades africanas nunca foram sistemas fechados imunes à integração de novos elementos e transformações ou mudanças resultantes de influência externa. Antes pelo contrário, elas mantêm estruturas abertas que num processo contínuo permitem que novas realidades e elementos vindos de fora possam ser absorvidos, transformados e adaptados aos contextos africanos (cf. p. 128-130).

Considerando estes aspectos, neste texto discute-se o caso do português de/em Moçambique, mostrando a sua inserção neste país, tendo em conta os desafios e perspectivas decorrentes das funções que lhe estão associados. A discussão será desenvolvida considerando primariamente a vasta literatura sobre a situação do português de/em Moçambique, que incide sobre a sua transformação simbólica e estrutural.

1.2. Pressupostos teóricos

A presença e preponderância de línguas de origem europeia nos países africanos tem sido matéria de ampla controvérsia, envolvendo todas as esferas da intelectualidade, incluindo escritores (NGUNGI WA THIONGO, 1987; MIA COUTO, 2009; OKARA, 1963), filósofos (APPIAH, 1997), políticos (OBOTE, 1967), académicos (LAITIN, 1992; DJITE, 2008; BOKAMBA, 2011, HEINE, 1992), entre outros. Segundo Schmied (1991) duas posições fundamentais se confrontam, ou seja, os abolicionistas e os adaptacionistas. Os abolicionistas mantêm uma posição crítica, associando a manutenção das línguas de origem europeia a efeitos negativos

como elitismo (MYERS-SCOTTON, 1993), desigualdades sociais, subdesenvolvimento (DJITÉ, 1991), alienação das massas (NGUNGI WA THIONGO, 1987), dependência (neo-)colonial (WOLFF, 2021), etc. Por outro lado, os adaptacionistas assumem uma posição pragmática, reconhecendo a dificuldade de se prescindir das línguas de origem colonial, face à situação multilingue prevalecente em muitos países africanos. Por exemplo, para Chinua Achebe, embora tenham uma origem colonial, essas línguas podem ser usadas, se houver utilidade nisso (cf. ALTOBBAI 2020: 9). Para além disso, apontam para o facto de elas estarem a passar por um processo de transformação, que as adaptam aos contextos dos países africanos (cf. NEW 1995: 308). Há variados estudos que apontam nesse sentido, como os que se debruçaram sobre processos de nativização, de que surgem as chamadas variedades não-nativas (KACHRU, 1982; 1986; SCHNEIDER, 2003; 2007). O processo de nativização envolve a transformação de uma variedade europeia, adoptando características derivadas do novo contexto onde a língua foi transplantada e é usada pelos locais. A nativização conduz à redefinição de padrões linguísticos, o que pressupõe a legitimação de novas variedades das línguas de origem europeia. Eventualmente, a legitimação é reforçada pela sua oficialização e nacionalização, ou seja, o seu reconhecimento como património e símbolo de uma nação e instrumento autorizado para uso em funções oficiais. A discussão desenvolvida neste texto confronta estes pressupostos com o caso do português em Moçambique. Assim, ela se debruça sobre a implantação do português em Moçambique, assumindo que esta língua está num processo de nativização.

2. Implantação do português em Moçambique

2.1. A fase pré-independência

A transplantação do português para Moçambique é uma consequência da colonização portuguesa, principalmente as políticas adoptadas após a implantação do Estado Novo, que preconizavam a ocupação e exploração efectivas das colónias, em harmonia com as directrizes e supostos “direitos” adquiridos com o Pacto de Berlim (cf. NEWITT, 1995: 415-6). Assim, as autoridades coloniais promoveram um conjunto de reformas administrativas e sócio-económicas, que incluíam uma legislação laboral para atrair investimento através da oferta de mão-de-obra barata (PENVENNE, 1995: 3) e a transferência da capital da colónia da Ilha de Moçambique para a cidade de Lourenço Marques, agora Maputo, onde se instalou uma máquina burocrática para apoiar o estado colonial e as actividades económicas, agora intensificadas pelos contactos com a África do Sul (NEWITT, 1995: 382). Assim, se criou um ambiente que permitiu a transplantação do português, não só porque era a língua da colonização, usada pelos colonos, como também porque se difundia entre elementos da população local, principalmente as elites locais integradas no aparelho político-administrativo e sócio-económico colonial (PENVENNE, 1992). Esta situação é testemunhada, por exemplo, por Penvenne (1992) ao indicar que a expansão do complexo portuário em Lourenço Marques, implantação da burocracia estatal e desenvolvimento do comércio internacional, não podia prescindir do uso de elementos da população local, dada a exiguidade de colonos portugueses. Portanto, não só havia necessidade de se dispor de comerciantes locais, artesãos e produtores de artigos de consumo, etc. como também de treinamento e educação baratos de mão-de-obra especializada que pudesse tornar as instituições burocráticas e comerciais

funcionais. Assim, nas primeiras décadas do século passado, já havia uma reduzida elite africana local com conhecimento do português, que, de facto, se distinguia e se impunha através de aptidões peculiares, que incluíam a escolarização em português e domínio da língua portuguesa. Estes indivíduos eram legalmente considerados como não-indígenas ou assimilados, o que significa que eram tidos como cidadãos portugueses.

Uma vez que a condição fundamental para a aquisição do estatuto de não-indígena ou assimilado era o domínio do português, a política assimilacionista estabeleceu um precedente: a emergência de uma ideologia social que relacionava mobilidade social no seio da população autóctone com o conhecimento de português. Por outras palavras, a língua portuguesa se tornou numa forma de capital social directamente ligada aos sistemas simbólicos, ideológicos e económicos que controlavam a mobilidade social e relacionavam as habilidades pessoais com a estratificação social dentro da colónia (FIRMINO, 1995).

Assim, durante o período anterior à independência, a aquisição do português pela população africana foi primariamente motivada pela posição que a língua mantinha nos sistemas ideológico, sócio-cultural e económico coloniais, ou seja, a sua associação com a ideologia colonial de “civilização”, a actividade institucional e a ascensão social. No entanto, as limitações impostas pelo sistema colonial, que não tinha nem a capacidade nem a vontade de alargar a classe de africanos “civilizados”, impediram a sua aquisição por um segmento amplo da população africana, de tal modo que quando se deu a independência de Moçambique, o português fazia parte do repertório linguístico de um grupo minoritário de moçambicanos, muitos dos quais concentrados nas zonas urbanas. Para muitos desses falantes, o português era uma língua-segunda, com prestígio, aprendida formalmente na escola e usada frequentemente nos domínios institucionais.

Como será demonstrado nos parágrafos seguintes, a associação do português com o prestígio e a ascensão social não desaparecem com a independência. Mais do que isso, este prestígio foi reforçado pelo facto de o português ter sido envolvido num novo quadro ideológico, que o promoveu e o adoptou como um importante símbolo da unidade nacional e como língua oficial de Moçambique. Como resultado, o uso e os utentes do português expandiram-se, como se demonstra posteriormente na Tabela 1.

2.2. A fase pós-independência

No Moçambique independente, para além de ser considerado símbolo de unidade nacional, o português foi adoptado como língua oficial, conforme consagrado na constituição. A escolha do português como língua oficial e da unidade nacional era/é previsível dada a história do seu uso em Moçambique, o tipo de diversidade linguística prevalecente no país (FIRMINO, 2011), as premissas ideológicas relacionadas com o tipo de sociedade concebida para o país, bem como a necessidade de co-optar as elites, que operavam em português, nas estruturas do poder e nas instituições burocráticas públicas. Com efeito, em consequência da política colonial e da ideologia linguística com que se correlacionava, o português surgiu como a língua de prestígio associada às actividades institucionais e à mobilidade social. Constituía também a marca simbólica fundamental da elite escolarizada, independentemente da origem étnica, regional ou racial.

Como o funcionamento das instituições nacionais do estado dependeu da co-optação da elite escolarizada, que operava em português, a oficialização desta língua foi

uma consequência lógica, reforçada pelo facto de nenhuma língua autóctone cobrir todo o território. O português, mesmo que não fosse falado por uma proporção significativa da população, ainda assim possuía falantes distribuídos por todo o território nacional, que o usavam sem conotar lealdades étnicas potencialmente divisivas. Deste modo, pode-se argumentar que a oficialização do português está ligada ao desenvolvimento de um quadro ideológico que associava o português à promoção da unidade nacional e à criação de uma consciência nacional, tal como já era evidente durante a luta armada anti-colonial para a libertação de Moçambique, quando o movimento nacionalista FRELIMO (Frente de Libertação de Moçambique) decidiu adoptar a língua portuguesa com a finalidade de preservar a unidade entre os moçambicanos de diferentes origens envolvidos na sublevação.

Sob o ponto de vista prático, assumiu-se também que a oficialização do português facilitava não só processos de interacção entre os cidadãos, como também o acesso ao conhecimento técnico-científico e a integração no mundo das nações. Tal visão é documentada na passagem seguinte da então Ministra da Educação e Cultura, extraída do discurso de abertura do I Seminário sobre o Ensino da Língua Portuguesa:

“A Língua Portuguesa é o meio de comunicação entre todos os moçambicanos que permite quebrar as barreiras criadas pelas línguas maternas. Através dela, a ideologia do Partido FRELIMO, que encarna os interesses das massas trabalhadoras e exprime os seus valores revolucionários, é difundida e estudada para ser aplicada, orientando o nosso Povo na luta pela criação de uma sociedade justa, próspera e feliz, a Sociedade Socialista. A Língua Portuguesa é também a língua veicular do conhecimento científico e técnico. Na alfabetização de milhares de trabalhadores, operários e camponeses, ela desempenha um papel importante pois fornece os instrumentos necessários para orientar e controlar a produção, em suma, para melhorar o bem-estar social e material. (...) É ainda utilizando a Língua Portuguesa que comunicamos com outros povos do mundo, transmitindo a rica experiência do nosso Povo e recebendo a contribuição do património cultural mundial” (TEMPO n.º. 471: 12).

3. A expansão do português

O facto de o português, a língua do inimigo, ter sido apropriada pela FRELIMO no início da luta de libertação, deu uma forte legitimidade à sua adopção no Moçambique independente, sendo, por isso, que a tal adopção não podia, assim, ser percebida, aos olhos do público, como sinal de uma nostalgia colonial. Como já se argumentou, “diferentemente de muitos países do continente, a situação da língua portuguesa não é a de uma herança incómoda com carácter provisório enquanto se não encontra uma língua ‘genuinamente’ africana. (...) É um projecto que visa anular todas as consequências da arbitrariedade do traçado geográfico do País, dar-lhe uma identidade nacional e uma consciência cultural, através do povo que nele habita” (ROSÁRIO, 1982: 64-5). No mesmo diapasão, Ferreira (1988) já notara que os países africanos de língua oficial portuguesa tinham feito mais para a expansão do português nos poucos anos de independência do que aquilo que as autoridades coloniais portuguesas fizeram durante o período colonial (cf. p. 38). Por essa razão, Mia Couto afirmou o seguinte:

“Na realidade, as autoridades moçambicanas não mudaram a sua política linguística e o português permanecia na sua condição de língua oficial e unificadora. Fala-se hoje mais português em Moçambique do que se falava na altura da Independência. O governo moçambicano fez mais pela língua portuguesa que os quinhentos anos de colonização. Mas não o fez por causa de um projecto chamado lusofonia. Nem o fez para demonstrar nada aos outros ou para lançar culpas ao antigo colonizador. Fê-lo pelo seu próprio interesse nacional, pela defesa da coesão interna, pela construção da sua própria interioridade” (COUTO, 2009: 192-3).

Na sequência da sua adopção como língua oficial e símbolo de unidade nacional, o português tornou-se um meio importante de comunicação, não somente em cenários institucionais mas também nas interações quotidianas urbanas em espaços públicos e privados, tais como restaurantes, ruas, mercados, ambientes caseiros, etc. Por outro lado, a expansão da rede escolar, ainda que não seja satisfatória, contribuiu imenso para o aumento de falantes da língua portuguesa. Esta política favorável ao português ditou a sua apropriação social, com a consequente expansão para novos falantes e novos domínios. À medida que as pessoas usavam o português de diferentes maneiras, ele começou a transcender o seu papel de mero instrumento político e administrativo para se tornar num instrumento comunicativo e simbólico fundamental, relacionado com a vitalidade do novo ambiente nacional em Moçambique. Consequentemente, o número de falantes do português também subiu significativamente, especialmente nos centros urbanos, conforme se demonstra na Tabela 1.

TABELA 1

Percentagem de falantes de português (LP) e de línguas bantu (LB), como Língua Materna (L1) e como Língua-Segunda (L2) em Moçambique

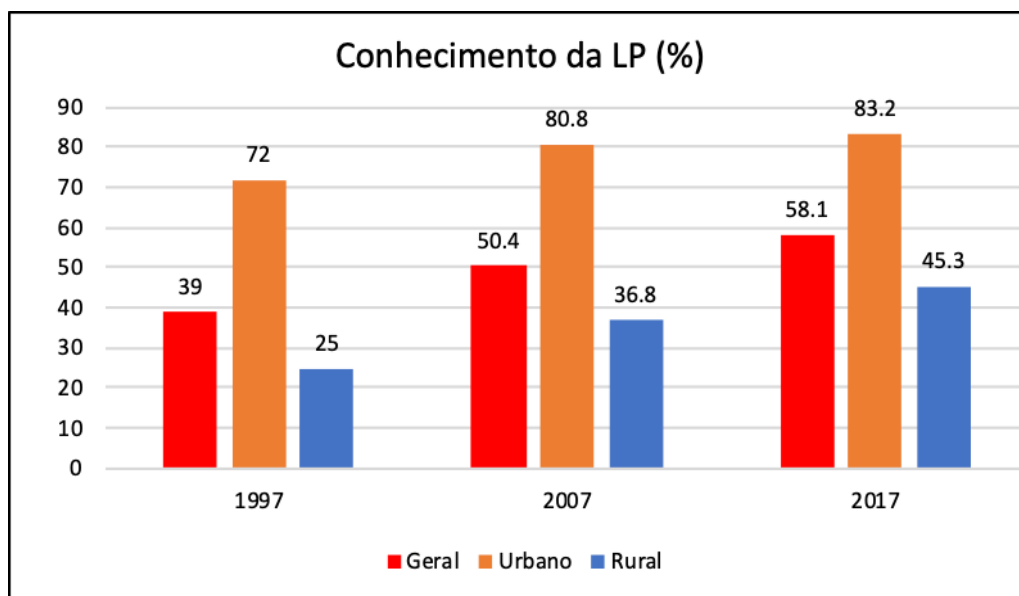
Ano	LP=L1	LP=L2 e LB=L1	LP=L1+L2	LB=L1 e LP=0
1980	1,2	23,2	24,4	75,6
1997	6,0	33,0	39,0	61,0
2007	10,4	39,9	50,3	49,7
2017	16,6	41,5	58,1	41,9

Fonte: Elaboração própria, com base em Conselho Coordenador do Recenseamento (1983), Chimbutane (2012), Firmino (2005), Firmino (2000) e INE (2019).

Com efeito, em 1980, segundo os dados do Recenseamento Geral de População (RGP) 24,4% da população de Moçambique reivindicou que sabia português, dos quais 23,2% como língua segunda e 1,2% como falantes nativos. Em 1997, segundo o Recenseamento Geral de População e de Habitação (RGPH), os que declararam que sabiam falar português já atingiam a cifra de 39,0%. Entre a população urbana esta cifra já rondava os 72,0%, contra 25,0% nas zonas rurais. Em 2007, segundo o Recenseamento Geral de População e de Habitação (RGPH), os que declararam que sabiam falar português já tinham atingido a cifra de 50,4%. Entre a população urbana esta cifra já rondava os 80,8%, contra 36,8% nas zonas rurais. Em 2017, segundo o Recenseamento Geral de População e de Habitação (RGPH), embora se tenha em conta que os dados estão em revisão, nota-se que os declararam que sabiam falar

português atingiam a cifra de 58.1%. Entre a população urbana esta cifra rondava os 83.2%, contra 45.3% nas zonas rurais (cf. Gráfico 1).

Gráfico 1. Falantes de português por zonas de residência



4. O português de/em Moçambique

4.1. Nativização

Do mesmo modo que o português no Moçambique independente está a adquirir novas funções sociais, está também a desenvolver características estruturais e retóricas típicas. O desenvolvimento dessas características é, contudo, uma continuação de um processo que começou antes da independência. Com efeito, mesmo antes da independência, o português falado em Moçambique incluía traços típicos largamente propagados, ou seja, os chamados moçambicanismos, usados até pelos colonos portugueses. Tais moçambicanismos podem ser exemplificados com elementos lexicais, como *milando* (equivalente a problema) *maningue* (equivalente a muito), *quinhenta* (cinquenta centavos, na forma europeia, também usada em Moçambique), *landim* (língua indígena falada em Maputo), *mainato* (empregado doméstico negro encarregado de lavar e engomar a roupa), entre outros.

Depois da independência, o uso do português alargou-se e os sinais da sua "moçambicanização" se expandiram. Enquanto o uso do português se alargava, os mecanismos que haviam contribuído para a aprendizagem e reforço do padrão linguístico se alteraram, dando origem à proliferação de novas formas linguísticas. Como indica Rosário (1982):

"O traumatismo do 'Pretoguês' foi desaparecendo e hoje qualquer cidadão faz questão de se exprimir correctamente em português e quantos deles, sendo responsáveis de sectores, não fazem brilhantes intervenções numa expressão recheada de neologismos de momento, estruturas totalmente novas e alheias à língua, interferências várias" (p. 65).

Albino Magaia, um jornalista e intelectual local bem conhecido, manteve uma opinião similar. Num dos seus artigos sobre este assunto, intitulado “Língua: Vida de um povo, espelho da sua história”, escreveu:

“... alterada a correlação política que ligava o português às línguas nacionais (ele transformou-se em língua oficial, língua de unidade) foi por isso e inevitavelmente desarmado do seu papel tradicional de instrumento de opressão cultural. Com uma criatividade exuberante as línguas nacionais violentam-no, viram-no e reviram-no, moldam-no por vezes à sua própria estrutura gramatical como podemos constatar no dia a dia da nossa revolução. Impossível pretender manter a pureza do português ou, o inverso pretender eliminar as influências por ela exercidas e consolidadas nas línguas nacionais”. (in TEMPO no. 477: 33)

4. 2. Índícios de nativização: alguns traços

As inovações que caracterizam a língua portuguesa em Moçambique abrangem múltiplas dimensões, como tem sido mencionado em vários estudos¹, que incluem aspectos fonético-fonológicos, morfo-sintáticos, semântico-pragmáticos e retóricos. Com base em vários estudos (DIAS 2002; 2009; GONÇALVES 2010; 2018; LOPES et al. 2002; 2016; MACHUNGO 2000, 2015; MENDES 2000; 2010; NHATUVE, 2017; TIMBANE 2013, entre outros), são indicados exemplos que ilustram alguns casos de propriedades gramaticais mais estáveis que indiciam a nativização do português em Moçambique.

4.2.1 Domínio fonético-fonológico

O domínio da fonética, fonologia e prosódia apesar de ainda não ser alvo de estudos mais consistentes, parece ser um dos que apresenta mais estabilidade, por ser fortemente marcado pelas línguas de substrato dos falantes, todas elas de origem bantu. É possível hoje, em Moçambique, sem grande margem de erro, identificar a proveniência regional dos falantes pela prosódia e pronúncia, em particular de falantes dos grandes grupos linguísticos existentes no norte, centro e sul do país.

Apresentamos de forma generalizada aqueles fenómenos que se considera serem mais relevantes e estáveis, no domínio do vocalismo e consonantismo; note-se que os casos descritos não são realizados por todos os falantes, dependendo em grande parte, como foi acima referido, das suas línguas de substrato².

¹ Bibliografia geral sobre o Português de Moçambique pode ser consultada no site da Cátedra de Português Língua Segunda e Estrangeira da Universidade Eduardo Mondlane em www.catedraportugues.uem.mz

² Para melhor compreensão de alguns fenómenos linguísticos, tomou-se como referência o português padrão europeu, por ser aquele que oficialmente foi adoptado no país.

i. Vocalismo

ADOPÇÃO DO PADRÃO SILÁBICO CV

Epêntese vocálica em sílabas fechadas por consoante

Palavra	PM	PE
culpado	[kulu'padu]	[kuɫ'padu]

Inserção de vogal em ataques ramificados

Palavra	PM	PE
grande	[ga'rãde]	['grãd]

Abertura de vogal em posição átona

Palavra	PM	PE
celeste	[sɛ'ɫɛjte]	[si'ɫɛt]

Nasalização/ Desnasalização de vogais

Palavra	PM	PE
Joaquim	[jõa'ki]	[jue'kĩ]

Não ditongação antes de consoante palatal

Palavra	PM	PE
Cerveja	[sɛr'veja]	[sir'vejʒe]

ii. Consonantismo

Desvozeamento de consoantes

Palavra	PM	PE
Bandeira	[pã'tɛra]	[bẽ'dejɾe]

Inconsistência no uso da vibrante simples e múltipla

Palavra	PM	PE
Muro	['muro]	['muru]

4.2.2. Domínio léxico-sintáctico e morfossintáctico

Estudos efectuados sobre aspectos morfossintácticos e semânticos do PM revelam que as inovações linguísticas afectam fundamentalmente preposições e verbos. Gonçalves (2018) acentua a necessidade de descrição desta variedade não-nativa do Português numa perspectiva que permita fazer generalizações do ponto de vista formal da gramática da língua. Deste modo refere que a mudança gradual de uma perspectiva externa, aprendizagem da língua, para uma perspectiva interna, aquisição da língua, ("internalized and externalized language" desenvolvida na teoria generativa de Chomsky 1986) tem-se revelado importante na análise de processos linguísticos em situação de contacto de línguas. Observem-se os exemplos abaixo (exemplos retirados de Gonçalves, 2018):

1. os pais escondem **os filhos** [aos filhos] a verdade
2. O detective **desconfiou um** [desconfiou de] indivíduo
3. Elogiam **a uma pessoa** [uma pessoa]
4. Chegou cedo **na escola** [à escola]
5. **Ameacei-lhe** [ameacei-o] com o meu apelido

De acordo com Gonçalves (2018), a análise das formações apresentadas de 1-5 numa perspectiva externa da língua, pode levar à conclusão de que os falantes do PM que têm como L1 uma língua bantu apresentam deficiências no processo de aprendizagem da língua. No entanto, estas frases devem ser analisadas à luz de uma perspectiva interna da língua, ou seja, uma perspectiva que tome em conta a importância da interferência do conhecimento da gramática da língua materna, na formação de novas gramáticas; com esta abordagem podem explicar-se fenómenos linguísticos aparentemente sem qualquer relação e assim demonstrar, no caso do PM, que este é um sistema linguístico coeso sendo possível explicar as inovações linguísticas específicas da gramática desta língua (GONÇALVES, 2018: 71). Assim, a pesquisa efectuada permitiu identificar e fazer generalizações gramaticais que afectam as propriedades de selecção categorial dos verbos e o papel semântico de verbos preposicionados atribuindo novas propriedades às preposições e verbos. Por exemplo, nota-se frequentemente a supressão de preposições que no PE regem complementos com a função gramatical de objecto indirecto e oblíquo:

6. ele entregou **o pai** a faca/ **ao pai**;
7. o porco desconfiou **o milhafre/do milhafre**)

Há também uma tendência para associar algumas preposições verbais a papéis semânticos distintos dos que existem no PE:

8. a. elogiam **a uma pessoa** quando morre/ **uma pessoa**
b. ela chega sempre tarde **na escola/ à escola**
c. separei **com os meus pais** muito cedo/ **dos meus pais**

No que concerne à morfossintaxe há uma tendência no PM para o uso do clítico dativo em detrimento da forma acusativa usada no PE:

9. o chapeiro **deixou-lhe** fora da paragem/ **deixou-a**

Existe igualmente uma oscilação e inconsistência no uso das formas de tratamento, em particular da 2ª pessoa:

10. cuide da **tua** saúde/ **sua**

4. 2. 3. Dimensão lexical

Neste texto dar-se-á ênfase a aspectos lexicais por ser uma dimensão que se destaca nos processos de nativização (SCHNEIDER 2007: 56) e por haver desenvolvimentos recentes na sistematização e caracterização de inovações lexicais no PM.

Como anteriormente referido, o PM, cujas formas linguísticas no país vão desde a norma considerada próxima do PE, usada sobretudo na comunicação escrita (documentos oficiais, instituições de ensino, meios de comunicação escrita) à "norma" do PM usada sobretudo na comunicação oral (escola, rádio, televisão, emprego), possui características que o distinguem de outras variedades do português.

Na sua dimensão lexical, grosso modo, pode-se afirmar que o português de Moçambique é composto por (i) palavras comuns a outras variedades do português, (ii) palavras que pertencendo ao léxico comum adquirem traços linguísticos particulares resultantes ou não da aplicação de processos morfológicos e (iii) empréstimos a outras línguas, com destaque para línguas bantu e inglês.

As palavras nos exemplos 11. abaixo foram submetidas a processos morfológicos de que resultou a aquisição de traços semânticos particulares:

11. *sala-sombra* "sombra de árvore que serve de sala de aulas"
sograria "casa dos sogros"
raspar cartão 'passar o cartão de crédito na POS'

No PM também se observam palavras que não sofrendo modificações de forma adquirem ou não novos traços semânticos e desencadeiam novas estruturas de subcategorização, como em:

12. *escolar*_{VERB} "frequentar a mesma escola que alguém" (ex: *escolámos juntos*)
*nascer*_{V. TRANSITIVO} "dar à luz" (ex: *ela nasceu dois gémeos*)
*estrutura*_N "autoridade" (ex: *as estruturas da saúde viajavam em executiva*)

Nota-se igualmente no PM a existência de palavras que provêm de empréstimos, em particular às línguas bantu:

13. *bacela* “agrado que se faz ao comprador oferecendo-lhe uma porção do produto adquirido ou outro produto” (do Xichangana *kubasela*)
sura “vinho de palma”(do Gitonga *surha*)
txovar “empurrar” (do verbo *Xichangana kuchova*)

Estas manifestações linguísticas impõem um tratamento ontológico do PM, o que passa pela sua normalização. A normalização de uma língua supõe a produção de instrumentos de normalização linguística contando-se de entre os mais importantes, as gramáticas, os dicionários, os tratados de ortografia e os prontuários ortográficos (DUARTE, 2000:27). A aceitação/adopção pela sociedade de estruturas de vários domínios linguísticos é um elemento fundamental na normalização linguística.

A inexistência em Moçambique destes instrumentos reguladores e a necessidade da sua construção, os diversos trabalhos de descrição lexical já disponíveis, criaram as condições para a elaboração de recursos lexicais que contribuirão para a normalização linguística no país. Destes, destacam-se:

- o Observatório de Neologismos do Português de Moçambique (ONPM) que tem como escopo de actuação a detecção, recolha e sistematização de formas neológicas que ocorrem no discurso quotidiano dos falantes moçambicanos do português, fundamentalmente em meios de comunicação escrita e oral ; o ONPM, em constante actualização, disponibiliza através da internet, uma base de dados lexicais com informação que permite descrever e analisar os recursos linguísticos de que os falantes se servem para actualizar o seu léxico; actualmente reúne mais de 1600 neologismos lexicais.

- o Vocabulário Ortográfico do Português de Moçambique (VOMOLP) criado no âmbito do VOC – Vocabulário Ortográfico Comum da Língua Portuguesa, instrumento que viabiliza uma gestão democrática inovadora e supranacional da língua portuguesa (IILP, 2011). O VOMOLP disponível e de acesso livre na internet, contém cerca de 40 mil entradas, com informação relativa às características formais do léxico do português de Moçambique, tais como a ortografia, a flexão, a classe gramatical; inclui também um vasto vocabulário onomástico (toponímico) e léxico de especialidade.

- o Dicionário do Português de Moçambique (DiPoMo), que beneficiando dos instrumentos teóricos e metodológicos e da experiência acumulada na elaboração do VOMOLP, visa disponibilizar, em formato digital e físico, um recurso linguístico de larga escala que descreve e caracteriza o léxico do PM e com efeitos multiplicativos, uma vez que poderá:

- permitir, através do seu uso em instituições de ensino, a disseminação e consolidação da língua, a nível interno e externo facilitando a adopção de políticas que favorecem o pluricentrismo linguístico;
- constituir-se como base para a construção de materiais didáticos adequados à realidade linguística do país;
- servir de base para a criação de recursos computacionais (ex. correctores ortográficos, indexação semântica).

Na concepção do DiPoMo adoptou-se uma perspectiva integralista, construindo um dicionário que descreve o léxico do PM como um todo funcional, não numa dimensão contrastiva que tem caracterizado muitos dos estudos que se têm produzido sobre o PM. Para se atingir este fim, optou-se por uma dupla via de obtenção de dados que tem em conta a situação linguística do país. Assim, (i) recorrer-se-á à tradição lexicográfica do português

(procedendo à mineração de dicionários gerias de língua para se obterem as acepções e definições marcadas nesses dicionários como características do PM) e (ii) uso de tecnologias modernas para a exploração de corpora digitais (e-lexicografia). Dar-se-á particular atenção ao tratamento de formas alógenas, em particular as oriundas das LB, bem como à toponímia com relevância administrativa.

Para a constituição do corpus textual e paralexigráfico e tendo em conta a variação regional do português no país, em particular no domínio lexical, estabeleceram-se pontos focais em cada uma das regiões administrativas/linguísticas do país, para garantir uma maior homogeneização dos dados. Para além das fontes escritas, estes têm a incumbência de registar palavras que, sendo de uso cotidiano, incluindo em meios de comunicação social, não estão atestadas na forma escrita.

Estando no seu primeiro ano de execução, os resultados preliminares da pesquisa conducente à elaboração do DiPoMo são ainda do domínio da construção de ferramentas (foi construída uma plataforma de interface para inserção de dados textuais – a Plataforma Timbila), selecção, obtenção e limpeza de dados textuais num total de 30 milhões de tokens, dos 60 milhões previstos, e a selecção e obtenção de fontes paralexigráficas primárias e secundárias.

Os pressupostos teóricos e metodológicos adoptados na constituição da base de dados do DiPoMo fazem dele também um recurso valioso para a elaboração das Terminologias Científicas e Técnicas.

5. Considerações finais

O presente artigo discutiu a apropriação da língua portuguesa em Moçambique, destacando que esta língua tem estado a desempenhar funções importantes no país, como língua oficial e símbolo da unidade nacional, transcendendo, deste modo, a sua anterior posição de língua colonial. Na sequência desta situação, o português está a sofrer um processo de nativização, já que se associa a novos valores sócio-simbólicos e novos traços linguísticos, tendo as línguas e culturas de substrato, as línguas e cultura bantu predominantes no país, um papel preponderante neste processo. No artigo, exemplificam-se alguns desses traços característicos do português de Moçambique, nos domínios fonético, fonológico e morfossintáctico tendo-se optado por destacar os aspectos lexicais por ser esta a dimensão que mais se evidencia nos processos de nativização.

6. Referências bibliográficas

ALTOBBAI, A. S. The Paradox of Cultural Decolonization through the Colonizer's Language in Achebe's *Things Fall Apart*. *International Journal on Studies in English Language and Literature* (IJSELL), Volume 8, Issue 5, pp. 8-12, 2020. Disponível em: <http://dx.doi.org/10.20431/2347-3134.0805002>.

APPIAH, K. A. *Na casa de meu pai: A África na filosofia da cultura*. Rio de Janeiro: Contraponto, 1997.

BOKAMBA, E. G. The Africanization of English. In: KACHRU, Braj. (org.). *The other tongue: English across cultures*. Urbana, IL: University of Illinois Press, 1982, p. 77–98.

BOKAMBA, E. G. Ukolonia in African language policies and practices. In Bokamba Eyamba G. et al. (eds.). *Selected proceedings of the 40th annual conference on African linguistics*, p. 146-167. Somerville, MA: Cascadilla Proceedings Project, 2011.

BOURHIS, R. *Language policies and language attitudes: Le monde de la Francophonie*. In: RYAN, E. & GILES, H. (org.). *Attitudes toward language variation*. London: Edward Arnold, 1982, p. 34-62.

CHIMBUTANE, F.S. *Panorama linguístico de Moçambique-Análise dos dados do III recenseamento geral da população e habitação de 2007*. Maputo: INE, 2012.

CONSELHO COORDENADOR DO RECEASEAMENTO. *Recenseamento geral da população: Informação pública*. Maputo: INE, 1983.

COUTO, M. *E se Obama fosse africano? e outras interinvenções*. Lisboa: Editorial Caminho, 2009.

DIAS, H. *Mini-Dicionário de Moçambicanismos*. Maputo: Edição da autora, 2002.

DIAS, H. (org.) *Português Moçambicano. Estudos e reflexões*. Maputo: Imprensa Universitária, 2009.

DJITÉ, P. G. *Langues et développement en Afrique. Language problems and language planning*, vol. 15, No. 2, pp. 121-138, 1991.

DJITÉ, P. G. *The Sociolinguistics of development in Africa*. Clevendon: Multilingual Matters, 2008.

DUARTE, I. *Língua Portuguesa: instrumentos de análise*. Lisboa: Universidade Aberta, 2000.

FÉRAL, C. Les «variétés» du français en Afrique. *Stigmatisations, dénominations, réification: à qui la faute? Cahiers de sociolinguistique*, CAIRN, n. 15, p. 41-53, 2010-1. Disponível em: <https://doi.org/10.3917/csl.1001.0041>. Acesso em: 21 Set 2020.

FERREIRA, M. *Que futuro para a língua portuguesa em África?* Linda-A-Velha: ALAC/Edições A Preto e Branco, 1988.

FIRMINO, G. *A Situação linguística de Moçambique: Dados do II Recenseamento Geral da População e Habitação de 1997*. Maputo: INE, 2000.

FIRMINO, G. *A “questão linguística” na África pós-colonial: o caso do português e das línguas autóctones em Moçambique*. Maputo: Texto Editores, 2005.

FIRMINO, G. *Nation-statehood and linguistic diversity in the postcolony: The case of Portuguese and indigenous languages in Mozambique*. In: ANCHIMBE, Eric A. & MFORTEH, Stephen A. (org.), *Postcolonial Linguistic Voices: Identity choices and representations*, Berlin/Boston: Walter de Gruyter, 2011, p. 99-117

FIRMINO, G. & GUISSIMO, M. *Situação Linguística de Moçambique*. Comunicação apresentada nas Jornadas Científicas da FLCS 2019, Universidade Eduardo Mondlane - Faculdade de Letras e Ciências Sociais, 17 e 18 de Outubro de 2019, 2019.

GANHÃO, F. *O papel da língua portuguesa em Moçambique*. Comunicação apresentada no I Seminário Nacional sobre o Ensino da Língua Portuguesa, Maputo, Moçambique, 1979.

GONÇALVES, P. *Génese do português de Moçambique*. Lisboa: Imprensa Nacional-Casa da Moeda, 2010.

GONÇALVES, P. *O português em África*. In: RAPOSO, Eduardo Buzaglo Paiva; NASCIMENTO, Maria Fernanda Bacelar do; MOTA, Maria Antónia Coelho da; SEGURA, Luísa; MENDES, Amália (org.). *Gramática do Português*. Lisboa: Fundação Calouste Gulbenkian, 2013, p. 157-178.

GONÇALVES, P. *A génese do Português de Moçambique*. Lisboa: Imprensa Nacional Casa da Moeda, 2010.

GONÇALVES, P. *Research on L2 varieties of European languages: From descriptive to formal grammars*. In: LOPES, L.;

GONÇALVES, P.; AVELAR, J. (org.) *Issues in Hispanic and Lusophone Linguistics 20. The Portuguese Language continuum in Africa and Brazil*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2018.

HEINE, B. *Language policies in Africa*. In: Herbert, R. K. (ed.). *Language and society in Africa*. Johannesburg: Witwatersrand University Press, pp. 23-35, 1992.

INE-INSTITUTO NACIONAL DE ESTATÍSTICA. *IV recenseamento geral da população e habitação 2017- Resultados definitivos*, Maputo: Instituto Nacional de Estatística, 2019.

KACHRU, B. B. *Models for non-native englishes*. In: KACHRU, Braj B. (org.). *The other tongue: English across cultures*. Urbana, IL: University of Illinois Press, 1982, p.31-57.

KACHRU, B. B. *The alchemy of English: The spread, functions and models of non-native Englishes*, Oxford: Pergamon Press, 1986.

LAITIN, D. *Language repertoires and state construction in Africa*. Cambridge: CUP, 1992

LIPSKI, J. M. *The Spanish of Equatorial Guinea. The dialect of Malabo and its implications for Spanish dialectology*. Series: Beihefte zur Zeitschrift für romanische Philologie, 209. Mouton: De Gruyter, 1985.

LIPSKI, J. M. *The Spanish language of Equatorial Guinea*. *Arizona Journal of Hispanic Cultural Studies*, v. 8, n.1, p. 115-130, Jan. 2004. Disponível em: <http://www.jstor.org/stable/20641705>. Acesso em: 20 Set 2020.

LOPES, A. et al. *Moçambicanismos: Para um léxico de usos do português moçambicano*. Maputo: Livraria Universitária, 2002.

LOPES, A. et al. *Com todos os Efes-e-erres: para um léxico de usos idiomáticos Português-Inglês*. Maputo: Livraria Universitária. 2016.

MACHUNGO, I. *Neologisms in Mozambican Portuguese*. Tese de Doutoramento – Legon-University of Ghana, 2000.

MACHUNGO, I. *Estratégias de criação lexical no português de Moçambique: aspectos da derivação sufixal*. *Revista Científica da UEM*, 1 (1), Edição Especial: 109-121. 2015.

MANNESSY, Gabriel. *De la subversion des langues importées: Le Français en Afrique Noire*. In: CHAUDENSON, Robert; ROBILLARD, Didier. (org.). *Langues, Économie et Développement*, (vol. 1). Institut d'Études Créoles et Francophones, Université de Provence, Diffusion: Didier Érudition, 1989, p. 133-145.

MENDES, I. *O léxico no português de Moçambique: aspectos neológicos e terminológicos*. Maputo: Promédia, 2000.

MENDES, I. *Da neologia ao dicionário: O caso do português de Moçambique*. Maputo: texto Editores, 2010.

MESTHRIE, Rajend/Bhatt; MOHA, Rakesh. *World Englishes: The study of new linguistic varieties*. Cambridge. Cambridge University Press, 2008.

MYERS-SCOTTON, Carol. *Elite closure as a powerful language strategy: the African case*. *International Journal of the Sociology of Language*, n. 103, p. 149-163, 1993. Disponível em DOI: <https://doi.org/10.1515/ijsl.1993.103.149>. Acesso 24 set 2020.

NEW, W.H. *New Language, New World*. In ASHCROFT, B., GRIFFITHS, G.& TIFFIN, H. (eds.) *The Post-colonial Studies Reader*, pp. 303-308, 1995.

NEWITT, Malyn. *A history of Mozambique*. Bloomington, Indianapolis: Indiana University Press, 1995.

NEWITT, Malyn. *A short history of Mozambique*, Johannesburg/Cape Town: Jonathan Ball Publishers, 2018.

NGUGI WA THIONGO. *Decolonizing the Mind: The Politics of Language in African Literature*. Harare: Zimbabwe Publishing House, 1987.

NHATUVE, D. *Reflexões sobre a normatização do Português de Moçambique*. Fórum linguístico, Florianópolis, v.14, n.2, 2017.

OKARA, G. *African speech ... English words*. Transition, Vol. 3, No. 10, pp.15-16, 1963.

PENVENNE, J. *The unmaking of an African bourgeoisie*, Lourenço Marques, Mozambique. Boston, MA: Boston University, African Studies Center (Working Papers No. 57), 1992.

PENVENNE, J. *African workers and colonial racism*. Portsmouth, NH: Heinemann, 1995.

OBOTE, M. *Language and national identification*. East Africa Journal, Vol. IV, No. 2, Nairobi: The East African Institute of Social and Cultural Affairs, pp. 3-7, 1967.

ROSÁRIO, L. *Língua portuguesa e cultura moçambicana: De instrumento de consciência e unidade nacional a veículo e expressão de identidade cultural*. Cadernos de Literatura. Coimbra. 12, p. 117-134, 1982.

SCHNEIDER, E.W. *The dynamics of New Englishes: From identity construction to dialect birth*. Language, vol. 79, n. 2, p. 232-281, 2003.

SCHNEIDER, E.W. *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press, 2007.

SCHMIED, J.. *English in Africa: An Introduction*. New York: Longman, 1991.

TENGAN, A. B. *European languages in African society and culture: A view on cultural authenticity*. in Putz, M. (ed.). *Language Contact and Language Conflict*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 125-138, 1994.

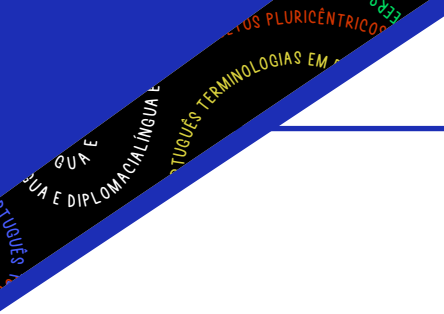
TIMBANE, A. *A variação e a mudança lexical da Língua Portuguesa em Moçambique*. Tese de Doutorado – Universidade Estadual Paulista “Júlio de Mesquita Filho”

Campus de Araraquara, São Paulo, 2013. Disponível em https://www.academia.edu/36943901/Tese_Doutorado_Alexandre_A_TIMBANE. Acesso em 15 Novembro 2021.

TIMBANE, A. *A lexicultura no PM*. *Linguagem: Estudos e Pesquisas*, 18 (2), 2016.

TUNDE-AWE, B. M. *Nativization of English language in a multilingual setting: The example of Nigeria*. Academic Journal of Interdisciplinary Studies, vol. 3, n. 6, p. 485-492, 2014 <https://www.mcser.org/journal/index.php/ajis/article/viewFile/4897/4739>, Disponível em: Doi:10.5901/ajis.2014.v3n6p485. Acesso em: 28 maio 2019.

WOLFF, E. (2021). *The ‘de-indigenization’ of African languages*. Academia Letters, Article 2702, 2021. Disponível em: doi.org/10.20935/AL2702. Acesso em: 15 Fevereiro 2022.



A importância das terminologias em português para a tradução e a interpretação na União Europeia

Joana Guimarães

Sobre o autor

Faculdade de Letras da Universidade do Porto, Professora Auxiliar, Doutoramento em Tradução. mguimaraes@letras.up.pt

Resumo

O multilinguismo, sendo um dos pilares sobre os quais assenta a construção do projeto europeu, comporta muitos desafios para todos aqueles que trabalham quer na tradução quer na interpretação da comunicação que diariamente ocorre no seio das instituições comunitárias. Sem um sólido trabalho de terminologia, levado a cabo por terminólogos em cooperação com os próprios tradutores e intérpretes e com os especialistas dos mais variados domínios, seria impossível levar a bom porto este projeto. São inúmeros os recursos terminológicos ao dispor dos profissionais das línguas, mas por vezes, sobretudo no que toca aos intérpretes (que é a realidade que a autora melhor conhece), estão muito dispersos e pouco alinhados uns pelos outros. Perante o crescente grau de especialização do discurso, nomeadamente no domínio jurídico - já de si um domínio com características muito peculiares - e perante mudanças nas condições de trabalho, como as que foram impostas pela pandemia da COVID 19, torna-se imperativo modernizar os instrumentos de trabalho. Outro desafio que os serviços linguísticos enfrentam e de que estão plenamente cientes é a crescente digitalização, que pode e deve ser encarada como uma oportunidade de reforçar, a nível político, a importância da língua portuguesa, em geral e a existência de terminologia de qualidade em português, em particular.

Palavras-chaves

Multilinguismo, terminologia, tradução, interpretação

1. Introdução

O multilinguismo é, sem sombra de dúvida, uma das maiores riquezas da diversidade cultural da União Europeia. Estamos perante um princípio que está consagrado no Artigo 3º do Tratado da União e que é defendido como um dos seus valores fundamentais. Só ele pode assegurar uma verdadeira transparência nos planos político, jurídico, económico e social, salvaguardar o pleno exercício da democracia e garantir a participação ativa dos cidadãos de cada Estado Membro na construção do projeto europeu. Mas, para que esta grande riqueza não se torne um entrave à criação de uma genuína integração europeia, é necessário, antes de mais, um trabalho rigoroso dos profissionais das línguas, os tradutores e os intérpretes, de modo a evitar transformar a União numa verdadeira Babel.

No que à língua portuguesa diz respeito, na União Europeia, ela é na maioria das vezes uma língua “não original”, ou seja, aquilo a que no âmbito dos Estudos de Tradução se chama uma língua de chegada, uma língua para a qual se traduz. Quer a tradução quer a interpretação são, na realidade, uma porta de entrada de terminologia na língua portuguesa: um sem número de novas realidades, novos conceitos, que obrigam à escolha de palavras para se poder comunicar a seu respeito. Paralelamente, a língua portuguesa vai sendo integrada, a par das restantes línguas oficiais da União, em memórias de tradução, bases de dados terminológicas ou memórias documentais naturalmente multilingues sobre os mais diversos temas, dos mais variados domínios do saber.

Numa tentativa de melhor explicitar a importância do trabalho terminológico no seio da União Europeia, tomemos como exemplo, por ser aquele que melhor conheço¹, a Direção Geral de Interpretação (SCIC) da Comissão Europeia em Bruxelas.

Desde os anos sessenta do século passado que existem terminólogos ao serviço das instituições europeias, certamente por se ter reconhecido a importância de que se reveste a terminologia para a qualidade da atividade dos vários grupos de trabalho e em última instância para a qualidade das decisões tomadas aos mais variados níveis. Os contributos para o sistema de terminologia da DG Interpretação provêm de mais do que uma fonte: dois terminólogos a tempo inteiro, dois chamados *rotateurs*, que são intérpretes que consagram um dia por semana ao trabalho de terminologia, e a rede *Terminologica*, que é constituída por dois representantes de cada cabine. Estes contribuem para a compilação de glossários e reúnem várias vezes por ano para debater questões terminológicas. Os glossários, por sua vez, baseiam-se em informações provenientes das várias reuniões em que os intérpretes são chamados a trabalhar e dos inúmeros documentos traduzidos que emanam da Direção Geral de Tradução e dos serviços de tradução das outras instituições. É ainda frequente os terminólogos da DG Interpretação organizarem cursos sobre o funcionamento do sistema de terminologia e a preparação de reuniões, destinados quer a intérpretes em início de carreira, quer aos mais experientes. Não que haja trabalho terminológico que possa ser feito com falta de rigor, mas a forma como o intérprete utiliza os recursos terminológicos reveste-se forçosamente de características específicas, dadas as circunstâncias em que exerce a sua profissão. Enquanto um tradutor dispõe de algum tempo para fazer a sua pesquisa

¹ Entre 1995 e 1997 trabalhei como agente temporária do então Serviço Comum Interpretação de Conferência (SCIC) da Comissão Europeia em Bruxelas e posteriormente várias vezes como intérprete freelance.

terminológica e se não ficar satisfeito à primeira, pode sempre tentar uma segunda consulta, o intérprete não dispõe desse tempo e precisa, por isso, de encontrar o que procura de forma mais rápida. Por vezes, a simples inclusão de uma fotografia numa base de dados pode contribuir para agilizar uma pesquisa, por exemplo, destinada a esclarecer uma dúvida do intérprete em pleno trabalho numa reunião em que se discutem cotas das diferentes espécies de peixe pescadas pelas frotas dos vários Estados Membros.

Atualmente, embora existam muitas ferramentas que podem contribuir para facilitar o trabalho de pesquisa terminológica quer de tradutores quer de intérpretes, a verdade é que é fácil apercebermo-nos de que a informação se encontra bastante dispersa e pulverizada por um número excessivo de bases de dados, de que são exemplo a *EUR-LEX*, base de dados da legislação da União Europeia, a base de textos legislativos do Conselho de Ministros (*COUNCIL*), a *ECB terminology*, a base de dados terminológica do Banco Central Europeu, a *EP Legislative Observatory*, a base de dados do Parlamento Europeu sobre decisões legislativas, o portal *EURAMIS* para pesquisa de documentos traduzidos de todas as instituições (que permite procurar em inúmeras combinações linguísticas), e a *LATE (Interactive Terminology for Europe)*, talvez a mais conhecida de todas as bases de dados terminológicas, já que tem uma versão de acesso público, para além daquela que se destina apenas a uso interno. Esta é uma base de dados que existe desde 1999 e cuja versão mais recente foi lançada em novembro de 2018; o sítio web deste projeto é administrado pelo Centro de Tradução dos Organismos da União Europeia, no Luxemburgo, e esta base conta já com mais de oito milhões de termos. Infelizmente, nem todas as entradas estão completamente preenchidas para todas as línguas, o que pode gerar alguma frustração a quem as consulta.

Especificamente destinado ao uso pelos seus intérpretes, os serviços da DG Interpretação desenvolveram um motor de pesquisa, o *LITHOS*, que permite aos intérpretes procurar termos em todos os glossários desta Direção Geral, com base em perfis predefinidos pelos próprios, nomeadamente levando em consideração a combinação linguística de cada um. É também possível escolher a língua de interface para uma utilização mais agradável e simples. A base de dados criada para o efeito é atualizada mensalmente, a fim de incorporar todas as novas entradas e quaisquer novos glossários. Para os intérpretes, o *LITHOS*, quando comparado com outros motores de pesquisa, como a *LATE* possui ainda a grande vantagem de poder ser usado quer *online* (nomeadamente a partir das páginas de preparação das reuniões da DG Interpretação) quer *offline*, já que o seu conteúdo pode ser totalmente descarregado, de modo a que os intérpretes possam consultar todos os glossários disponíveis da DG Interpretação. Esta possibilidade é particularmente útil, quando o intérprete é obrigado a trabalhar fora da cabine, por exemplo, em missões de acompanhamento de peritos nos vários Estados Membros. Como em tudo, nem sempre mais é necessariamente melhor e o facto de existirem tantas bases de dados, numa grande dispersão, não contribui certamente para a eficácia da pesquisa e comporta o risco da falta de uniformidade e concordância nas soluções apresentadas em cada uma delas.

Uma particularidade importante do trabalho de terminologia no âmbito da União Europeia prende-se com o facto de esta ser uma instituição que se ocupa dos mais variados assuntos, desde a agricultura e as pescas à propriedade intelectual, passando pela cooperação policial, os transportes ou as energias renováveis. De um modo geral, as outras organizações internacionais tendem a dedicar-se apenas ou predominantemente a um domínio específico. Esta característica implica que se tenha de lidar com questões terminológicas relacionadas com as mais variadas áreas do saber e que, para resolver algumas dessas questões, se tenha

de recorrer a peritos das mais diversas especialidades também. Embora seja cada vez mais fácil, para esclarecer dúvidas pontuais, encontrar informação de fontes fidedignas disponível na Internet, não deixa de ser necessário o recurso aos especialistas, para esclarecer questões de fundo. CORREIA; GONZÁLEZ (2003:304) falam da importância de estabelecer «redes formais ou informais, permanentemente actualizadas, de fontes, especialistas e divulgadores [...] para troca de informação terminológica». Estes especialistas podem estar nas comissões técnicas de normalização, nos ministérios, na representação permanente de cada Estado Membro junto da União Europeia (REPER), mas também nas universidades e respetivos centros de investigação, ou até mesmo nas empresas privadas. O seu contributo é inestimável, sobretudo para o trabalho de validação da terminologia.

O terminólogo é assim frequentemente chamado a desempenhar o papel de mediador entre o especialista e os tradutores/intérpretes, sendo que, por vezes, no caso dos intérpretes, eles próprios assumem a iniciativa de estabelecer contacto com o especialista (muitas vezes o cliente para quem trabalham em reunião) em busca da solução para as suas dificuldades.

São inúmeros os desafios que se colocam a quem se dedica ao trabalho terminológico nas instituições europeias. Cada vez é mais frequente a língua do texto/discurso original não ser a língua do autor, ou seja, quem escreve ou fala nem sempre o faz naquela que é a sua língua materna, o que pode refletir-se na qualidade da informação produzida. Assim, surgem por vezes, no âmbito da tradução, problemas de compreensão do texto original, decorrentes de interferências linguísticas. Estas interferências podem mesmo levantar falsos problemas terminológicos.

Acrescenta-se ainda uma outra exigência que se prende com a necessidade de o tradutor, ao traduzir um documento, dever utilizar a terminologia que foi usada em documentos anteriores no mesmo domínio, para assegurar a existência de coerência terminológica. MATOS (2007) fala num “lastro” terminológico que pode constituir um obstáculo, sempre que a solução encontrada no passado não é de todo a mais adequada. Um outro desafio decorre da existência de múltiplas fontes de referência que frequentemente não são coerentes entre si, como por exemplo, duas diretivas sobre o mesmo assunto, i.e., dois documentos com o mesmo estatuto, sendo que cada um faz uso de uma solução terminológica distinta.

Além disso, há domínios do saber em que não existem termos em língua portuguesa. MATOS (2007) dá o exemplo das centrais nucleares. Portugal não tem centrais nucleares, daí que grande parte da terminologia nesse domínio tenha sido criada por “importação”, nomeadamente no que diz respeito aos resíduos radioativos e ao seu tratamento². Outro exemplo curioso, referido pelo mesmo autor, é o da navegação fluvial, que, apesar de existir em Portugal, não atinge entre nós a complexidade e o desenvolvimento de outros Estados Membros, nomeadamente a Alemanha. Daí que se tenha sentido, no seio da DG Tradução a necessidade de criar de raiz uma grande parte da terminologia relacionada com a segurança e a sinalização deste tipo de navegação.

Acresce ainda que as línguas não são estáticas, são organismos vivos e vão sendo moldadas pelo uso que delas se faz, o que leva a que as soluções terminológicas sejam muitas vezes provisórias. Acontece por isso com relativa frequência ser necessário atualizar as terminologias por que se optou inicialmente. Ainda no âmbito da terminologia relativa às centrais nucleares e de acordo com MATOS (2007), por exemplo, «decommissioning»

² A propósito deste termo “resíduos”, MENDES DA COSTA (2003:311) chama a atenção, de uma forma muito interessante, para o modo como o termo “desperdícios” foi substituído, na linguagem administrativa e na legislação portuguesa pelo termo mais recente “resíduos”.

começou por ser «desativação», passou a «desclassificação» e hoje é «desmantelamento»; são assim criadas soluções terminológicas provisórias, que depois vão sendo (ou não) confirmadas e tornadas definitivas.

Nos últimos anos, a tradução e a interpretação jurídicas assumiram um papel de maior destaque no seio das instituições europeias, em grande parte devido ao aprofundamento da cooperação em matéria civil e penal decorrente do Tratado de Lisboa. Multiplicaram-se os grupos de trabalho em que participam delegações com uma formação jurídica cada vez mais especializada e daí advém a necessidade de traduzir os documentos que servem de base às reuniões desses grupos e de interpretar as intervenções dos membros dessas delegações. RAMOS (2021) enumera vários grupos, cujos nomes nos permitem já adivinhar o grau de especialidade da terminologia usada pelos participantes nas suas reuniões de trabalho: "Insolvência", "Sucessões", "Cessão de Créditos", "Direito Contratual Europeu", "Direito Europeu Comum de Compra e Venda", "Obtenção de provas em matéria civil e comercial", "Reconhecimento e execução de decisões em matéria civil e comercial".

Ainda no contexto da tradução/interpretação no domínio do Direito, importa recordar que o Direito Comunitário tem a particularidade de ser aquilo que AJANI; ROSSI (2006: 80) chamam a expressão multilingue de uma única mensagem («the multilingual expression of a single message»). Em cada sistema jurídico individual de cada um dos Estados Membros coexistem duas linguagens jurídicas diferentes: uma nacional e uma comunitária. Dito por outras palavras, desta vez inspiradas nas da intérprete polaca STEFANIAK (2017), a União Europeia tem o seu sistema jurídico próprio, apesar de não ter a sua própria língua, tendo por isso de "pedir emprestada" aos Estados Membros a terminologia das suas linguagens jurídicas. Na União Europeia, não se produzem vinte e quatro atos jurídicos, mas apenas um ato jurídico em vinte e quatro versões linguísticas diferentes. Daí advém a necessidade de haver coerência entre estas versões; essa coerência é muitas vezes apelidada de concordância multilingue (multilingual concordance). Mas a complexidade de todo este sistema vai ainda mais longe: se cada país tem o seu próprio sistema jurídico, que se distingue claramente de todos os outros, não se pode partir da premissa de que a cada língua corresponde um sistema jurídico: Áustria e Alemanha, por exemplo, partilham a língua, mas têm sistemas completamente diferentes. Na Bélgica há um único sistema jurídico e três línguas oficiais. Além desta singularidade que representa esta relação não exclusiva entre língua e sistema jurídico, a linguagem jurídica tem outras características peculiares: o facto de ser prescritiva, normativa exige-lhe um grau de clareza e precisão que não pode ser hipotecado em prol de qualquer outro argumento, mesmo sabendo nós que há diferenças entre um regulamento e uma diretiva no que toca ao seu carácter vinculativo: enquanto um regulamento é um ato legislativo vinculativo, aplicável em todos os seus elementos em todos os Estados Membros da União, uma diretiva é um ato legislativo que fixa um objetivo geral, que todos os países da UE devem alcançar, cabendo contudo a cada país elaborar a sua própria legislação para dar cumprimento a esse objetivo. Já uma decisão só é vinculativa para os seus destinatários, que podem ser um grupo de países, um país apenas ou até mesmo uma empresa dentro de um país. Todos são documentos que "descolam" de uma realidade jurídica comunitária e "aterram" numa realidade jurídica nacional, com tudo o que isso implica.

Mas a equação não ficaria completa, se não introduzíssemos uma nova variável, a diversidade de destinatários que, no caso concreto da União Europeia, podem ser políticos, especialistas, mas também cidadãos comuns que queiram, por exemplo, candidatar-se a algum tipo de financiamento e que são colocados por isso numa relação de concorrência

direta com os seus congéneres oriundos de outros Estados Membros, que receberam outra versão linguística do mesmo documento de candidatura.

Voltando à importância dos especialistas e ao papel de mediador do tradutor/intérprete: este contacto permanente entre uns e outros pode evitar que se instalem no âmbito da União variantes terminológicas erradas de termos há muito consagrados nas línguas nacionais. RAMOS (2021) refere um exemplo ilustrativo desta situação e que se prende com a escolha do termo português *confisco*. Esta escolha representa uma colagem excessiva ao francês, ao espanhol ou ao inglês “*confiscation*”; o termo em português tem uma conotação histórica datada, relacionada com um período muito concreto da História de Portugal e por isso não é um termo isento... Apesar de o termo *confisco* ter sido o que foi inicialmente escolhido e surgiu nos documentos traduzidos, a expressão correta é “*perda*” (de bens a favor do Estado) – ou seja, opõe-se a tradução oficial comunitária à expressão efetivamente usada no país. Uma conversa com uma profissional da área do Direito, membro da delegação portuguesa num dos grupos de trabalho sobre esta matéria terá, segundo RAMOS (2021), sido suficiente para dissipar dúvidas e esclarecer esta questão terminológica.

Há uma crítica que é frequentemente dirigida a tradutores e intérpretes que trabalham nas instituições europeias e que é a de que o facto de estes serem normalmente emigrados a residir longe de Portugal e de estarem em constante convívio com outras línguas constitui uma ameaça à qualidade e à vernaculidade do seu português. Cito CORREIA (2007) na revista *a folha*, para defender a injustiça desta acusação:

«Pela sua própria condição de expatriados, os profissionais da língua dignos desse nome redobram de apego aos valores da integridade linguística e da idiomaticidade, precavendo-se contra os efeitos indesejados da imersão quotidiana numa ou em várias línguas estrangeiras.»

Nos tempos que correm, o acesso a fontes de informação em português está muito facilitado: basta haver vontade e há uma vasta oferta de produtos audiovisuais disponíveis; as plataformas de *streaming* são apenas uma entre muitas possibilidades de manter o contacto com a língua portuguesa mais atual. Se aos mais céticos ainda restarem dúvidas quanto à tal “integridade linguística”, mediante a utilização de filtros adequados nos motores de busca, é possível, hoje em dia, fazer aquilo a que CORREIA (2007) chama a contraprova, ou seja, obter um valor aproximado, correspondente à frequência com que um determinado termo ocorre em textos gerados em Portugal e, por exemplo, nas instituições europeias. CORREIA (2007) sugere que se faça o teste com um termo, que considera um verdadeiro barbarismo, o termo “*plafonamento*”. Eu própria segui esta sugestão e verifiquei que os resultados são elucidativos: contrariamente ao que seria de esperar, há mais entradas deste termo nos textos gerados em Portugal do que nos outros...

O crescente grau de especialização tanto dos textos escritos quanto dos discursos proferidos, que os profissionais das línguas têm de traduzir e interpretar, obriga a uma constante atualização dos recursos disponíveis e à criação de novas ferramentas de trabalho que facilitem o desempenho das suas funções e contribuam para a qualidade dos serviços prestados. Mas esta necessidade torna-se ainda mais premente, se atendermos à velocidade a que se tem dado a chamada transformação digital. As tecnologias de processamento da linguagem natural (PLN) não constituem uma ameaça para os tradutores e os intérpretes e não deverão ser encaradas com desconfiança, pois delas se pode esperar um contributo significativo para a melhoria da qualidade do trabalho de uns e de outros.

No Plano Estratégico da Direção Geral de Interpretação para o período 2022-24 (**Strategic Plan 2020-2024 DG SCIC**), está inscrita a intenção de lançar o projeto *Interpreters' Digital Toolbox*, com o objetivo de modernizar a gestão da terminologia e da documentação, fornecendo aos intérpretes uma interface digital integrada, para lhes facilitar a preparação das reuniões. A extração de terminologia com base em combinações linguísticas específicas é apenas uma das possibilidades que se pretende por em prática. Além da preparação das reuniões, pretende-se também facilitar/agilizar o trabalho de pesquisa em cabine, através do acesso simplificado aos documentos pertinentes para cada reunião, bem como a *links* para bases de dados terminológicas e glossários. A longo prazo, existe inclusivamente o desejo de ir acrescentando novas funcionalidades, baseadas na tecnologia de reconhecimento de voz. À semelhança do que já acontece em relação ao domínio da tradução, em que se recorre ao termo *augmented translation*, para designar a integração no ambiente de trabalho de tradução (de modo a não implicar a saída deste ambiente) de recursos de informação externos, tais como pesquisa em base de dados terminológicas, ferramentas de tradução automática, etc., também neste mesmo documento da DG Interpretação recorre, por analogia, ao conceito de *augmented interpreter*, para fazer referência a ajudas digitais que facilitam a preparação de tarefas de interpretação e constituem um apoio em tempo real enquanto o intérprete está a trabalhar em cabine.

A crise desencadeada pela pandemia da COVID-19 acelerou a transformação digital e trouxe novos desafios, mas também proporcionou novas oportunidades. Integrar o multilinguismo nas reuniões em ambiente remoto ou em modo híbrido nem sempre se revela fácil, mas é fundamental para salvaguardar os valores da União. Lidar com o multilinguismo sempre foi, e continua a ser, aprender a viver dividido entre a utopia e a realidade! Ora, no advento da era digital, a União Europeia reconheceu como uma das suas prioridades políticas a Estratégia do Mercado Único Digital (2015). Esta aposta política tem implicado e continuará a obrigar a tomadas de posição claras relativamente à questão do multilinguismo.

Neste contexto, o Parlamento Europeu, no relatório «Language equality in the digital age -Towards a Human Language Project», de 2017, alerta para o perigo de a língua inglesa se tornar a “língua franca” desse Mercado Único Digital; se tal acontecesse, 60% da população da União Europeia seria deixada para trás, havendo grandes disparidades entre os vários Estados Membros. A União não sobreviverá se passar a ser um projeto de uma elite minoritária que comunica numa língua franca e se ignorar os desejos e desígnios dos seus fundadores. Há ainda uma outra opção que por vezes se coloca sobre a mesa relativamente a esta matéria: um cenário em que as seis línguas mais faladas na União sejam as línguas de comunicação no âmbito do Mercado Único Digital. Tal solução excluiria 14% da população, o que equivale a mais de 70 milhões de cidadãos europeus. Nesse cenário, haveria onze Estados Membros em que mais de 50% da população ficaria impossibilitada de usufruir das vantagens e potencialidades do Mercado Único Digital. É naturalmente que esses cidadãos são aqueles que já por si se encontram em posição mais vulnerável, ou seja, os idosos, as pessoas com menos qualificações, aquelas que possuem rendimentos mais baixos e as que são oriundas de meios desfavorecidos. O multilinguismo é uma bandeira da diversidade cultural, mas é também uma pedra angular da democracia participativa; é através do multilinguismo que a União Europeia se aproxima dos cidadãos, pois estes veem que a União fala a sua língua.

É por isso que é tão importante que a crescente digitalização da sociedade europeia não abra um fosso entre as línguas mais utilizadas e as menos utilizadas. É uma realidade que as tecnologias da linguagem, presentes nos motores de pesquisa, nas redes sociais

ou nos serviços de comércio eletrónico são primordialmente disponibilizadas em inglês (apesar de existirem fabricantes, tanto europeus como de outros cantos do mundo, que desenvolvem estas tecnologias também para outras línguas europeias como o alemão, o francês ou o espanhol). Seria, contudo, extremamente importante que os conteúdos fossem disponibilizados em mais línguas e que a União Europeia, através das suas opções políticas e de financiamento (nomeadamente da investigação) se empenhasse no desenvolvimento de tecnologias da linguagem nas línguas oficiais que são menos faladas, como é o caso do português. A União e os seus Estados Membros partilham a grande responsabilidade de, a nível nacional, nas instituições públicas, mas também a nível local, por exemplo nas Universidades, contribuir para a preservação de todas as suas línguas no mundo digital.

O objetivo primeiro de um trabalho terminológico sólido e constante é sempre o de aumentar a eficácia da comunicação, mas não esqueçamos que, além da sua função comunicacional, a língua tem uma forte componente identitária. Logo, se houver terminologias em português de qualidade ao dispor dos serviços linguísticos da União Europeia, o serviço prestado a Portugal será melhor, facilitar-se-á o trabalho nas várias reuniões de nível técnico, dar-se-á rigor ao discurso e confiança a quem o profere e um empurrão à vontade política de defender a língua portuguesa e lhe atribuir o lugar que por direito deve ocupar neste grémio multilingue e multicultural.

6. Referências bibliográficas

AJANI, G.; ROSSI, P. *Multilingualism and the Coherence of European Private Law*. In: POZZO, B.; JACOMETTI, V. (Hrsg.). *Multilingualism and the Harmonisation of European Law*. Alpen aan den Rijn: Kluwer Law International, 2006, pp 79-91.

COMISSÃO EUROPEIA. *Strategic Plan 2020-2024 DG Interpretation*: https://ec.europa.eu/info/system/files/scic_sp_2020-2024_en.pdf

CORREIA, P.; GONZÁLEZ, L. *A propósito da relação entre terminólogo e especialistas. O caso do SdT da Comissão Europeia*. In: CORREIA, M. (org.): *Terminologia e indústrias da língua*. Actas do VII Simpósio Ibero-Americano de Terminologia. Lisboa: ILTEC, 2003, pp 287-307.

CORREIA, R. *A língua portuguesa vinte e um anos depois. a folha - Boletim da língua portuguesa nas instituições europeias*, 25 (Número especial), 2007. Disponível em: <http://ec.europa.eu/translation/portuguese/magazine>. Acesso em nov. 2021.

MATOS, C. *A tradução de textos científicos e tecnológicos na Comissão Europeia. a folha - Boletim da língua portuguesa nas instituições europeias*, 25 (Número especial), 2007. Disponível em: <http://ec.europa.eu/translation/portuguese/magazine>. Acesso em nov. 2021.

MENDES DA COSTA, A. *Terminologia política e política terminológica*. In: CORREIA, M. (org.): *Terminologia e indústrias da língua*. Actas do VII Simpósio Ibero-Americano de Terminologia. Lisboa: ILTEC, 2003, pp 309-321.

PARLAMENTO EUROPEU. *Language equality in the digital age - Towards a human language project*. Study: <https://op.europa.eu/en/publication-detail/-/publication/fa0a50e7-cda4-11e7-a5d5-01aa75ed71a1/language-en>

RAMOS, I. A interpretação jurídica na União Europeia: promover a unidade na diversidade. Comunicação em Conferência. 14. Deutscher Lusitanistentag, Universität Leipzig, 2021.

STEFANIAK, K. *Terminology work in the European Commission: Ensuring high-quality translation in a multilingual environment*. In: SVOBODA, T; BIEL, L.; LOBODA, K. (eds.). *Quality aspects in institutional translation*. Berlin: Language Science Press, 2017, pp. 109-121.

A língua portuguesa no espaço dos países africanos de língua oficial portuguesa da variedade do português de São Tomé e Príncipe à definição de um padrão

Abigail Tiny Cosme

Sobre o autor

Faculdade de Ciências e Tecnologias da Universidade de São Tomé e Príncipe. Professora do Departamento de Língua Portuguesa. Mestre em Linguística.
abigail.tiny@hotmail.com

Resumo

São Tomé e Príncipe é um país insular constituído por duas ilhas principais, totalizando cerca de 200 mil habitantes. Trata-se do país com menor número de população da Comunidade dos Países de Língua Portuguesa (CPLP) e o mais pequeno, em área, de África. Não obstante a dimensão, este caracteriza-se por uma grande diversidade linguística e cultural, resultado de um contexto histórico complexo.

Um conjunto de línguas compõe o mosaico linguístico do país, nomeadamente a Língua Portuguesa, que possui o estatuto de língua oficial, assim como o Forro, o Lung'le, o Angolar, o crioulo de Cabo Verde, o Português dos Tongas e resquícios de línguas do grupo Bantu.

O processo de formação das línguas crioulas nacionais está estreitamente relacionado com a história e a evolução do país, tendo cada estágio contribuído para a criação de condições favoráveis ao seu surgimento. A primeira fase, denominada da fase de habitação, iniciada ainda no século XV, por meio dos escravizados provenientes do delta do Níger, particularmente do antigo Reino de Benim (Hagemeijer 2011), sendo estes, sobretudo, falantes do Edo, uma língua da família Edóide. Na segunda fase, de plantação, o resgate dos escravizados foi transferido para zonas Bantu, com enfoque para o Reino do Congo, sendo estes falantes de Línguas Bantu, tais como o Kikongo e o Kimbundo.

No período colonial, o Português era considerado a língua das elites, da alfabetização e do acesso ao poder. Após a independência, em 1975, manteve o estatuto privilegiado, tendo a massificação do ensino, bem como a possibilidade de maior circulação e fixação das populações conduzido a um aumento substancial do número de falantes.

Neste artigo far-se-á uma análise à situação linguística, procurando aflorar as idiosincrasias do Português de e em São Tomé e Príncipe, fundamentais para um olhar crítico sobre o surgimento de uma variedade nacional, bem como da respetiva norma- padrão.

Palavras-chaves

Variedade; Norma-padrão; Português popular; Línguas crioulas.

1. O Contexto Linguístico em São Tomé e Príncipe

1.1. A origem dos Crioulos do Golfo da Guiné

Os dados históricos indicam que as ilhas de São Tomé e Príncipe foram “descobertas” nos anos de 1471/2 (Caldeira, 2006) e a de Ano Bom entre 1483 e 1501 (Caldeira, 2010). A primeira a ser povoada de forma permanente foi a de São Tomé, em 1493, sendo esta a ilha maior e centro das atividades económicas no Golfo da Guiné. A ilha do Príncipe terá sido povoada por volta do ano 1500 aquando da sua doação, por decreto real. Os Angolares têm sido descritos como descendentes de escravos quilombolas originários do primeiro povoamento de São Tomé (Caldeira, 2004; Ferraz, 1974; Lorenzino, 1998; Seibert, 2007). As fugas de escravizados estão amplamente documentadas desde as fases iniciais da habitação e tornaram-se uma séria ameaça para a povoação e as plantações a partir de 1530.

A ocupação de São Tomé pode ser dividida em duas fases: a fase de habitação e a fase de plantação. A primeira fase corresponde ao período entre 1485, data do primeiro povoamento, e cerca de 1515, altura que marca o início do ciclo do açúcar. A segunda fase vai até ao final do século XVI (Sousa, 1990). A história da primeira fase está diretamente relacionada com o comércio de escravizados no delta do Níger e particularmente com o antigo Reino de Benim, situado na atual Nigéria (Hagemeijer 2011). Há um consenso generalizado de que o número de pessoas escravizadas de origem Bantu aumentou significativamente depois de 1520, período que correspondeu ao rápido declínio do comércio do delta do Níger. O contacto entre os portugueses e os escravizados resgatados do delta do Níger, sobretudo do antigo Reino de Benim, na fase de habitação da ilha, criou as condições ideais para a formação de um pidgin. Neste período, havia um contacto intenso e direto entre os portugueses e os escravizados africanos do delta do Níger, que seriam sobretudo falantes do Edo, uma língua da família Edóide. Este pidgin nativizou entre as gerações seguintes de escravos, tendo dado origem ao que se designa por proto-crioulo do Golfo da Guiné (PCGG) (Hagemeijer 2009, 2011). Esta língua terá ramificado no tempo e no espaço, dando origem aos quatro CGG: Forro, Angolar, Lung’le e Fa d’Ambô. Há um conjunto de estudos que dão conta da relação de proximidade existente entre os CGG. Durante muito tempo, o ANG foi um crioulo relativamente desconhecido e associado às línguas Bantu.

Do ponto de vista linguístico, as línguas do delta do Níger da primeira fase tiveram visível impacto tipológico enquanto estrato africano primário na formação do PCGG. As estratégias de serialização verbal, de reflexivização com a palavra para corpo, a existência de implosivas e de lábio-velares são alguns dos exemplos que demonstram esse impacto. A fase de plantação, com a chegada maciça de escravos Bantu em detrimento de escravos do delta do Níger, contribuiu para a afirmação do papel do Bantu ocidental, nomeadamente

do Kikongo e do Kimbundu, nestes crioulos, com especial incidência a nível do léxico e da fonologia (Hagemeijer 2009).

As claras diferenças que hoje se podem observar nos quatro CGG são resultado do relativo isolamento em diferentes momentos da história, do efeito de contato posterior e do papel da mudança interna.

Em suma, grande parte dos estudiosos sobre esta área defende, generalizadamente, que os quatro CGG partilham uma origem comum (Hagemeijer 1999, 2009; Lorenzino 1998; Maurer 1995, 1999; Schang 2000).

Quanto ao Português, esteve sempre presente nas fases de habitação e plantação, considerando a presença dos colonos vindos da antiga metrópole, sendo esta a língua da elite colonial, de pouco acesso à restante população. Sobre ela abordaremos mais adiante.

2. Outras línguas

Ao longo da história recente das ilhas de São Tomé e Príncipe (STP), línguas têm ocupado espaços e lugares conducentes com as dinâmicas socioculturais, políticas e conjunturais. A partir de meados do século XIX, em consequência da introdução da cultura do cacau e do café, registou-se um aumento substancial da população com a chegada dos contratados provenientes de diferentes origens, nomeadamente Cabo Verde, Angola e Moçambique, dos então territórios ultramarinos, assim como de Benim, Gabão, Camarões, Serra Leoa e da Libéria. Com eles trouxeram, também, as suas línguas de origem, tais como o crioulo de Cabo Verde e dialetos do Kimbundo e do Umbundo. Considerando que os contratados de Cabo Verde deslocavam-se mais frequentemente em família e, considerando, também, a relativa homogeneidade da língua, criaram-se as condições favoráveis para que o crioulo fosse preservado e transmitido entre as gerações até aos dias de hoje. O crioulo de Cabo Verde, ou Kaboverdianu, continua a ser uma língua dominante em variadas roças, estando amplamente consolidado na ilha do Príncipe. Em sentido contrário, os falantes do Kimbundo e do Umbundo não tiveram as mesmas condições, registando-se maior número de casos de repatriamento de serviços angolanos e moçambicanos (HAGEMEIJER, 2009a). No entanto, em contexto de predominância destas línguas, o contacto com o português resultou no surgimento do Português dos Tongas (BAXTER, 2002, 2004), referindo-se o termo Tonga a descendentes dos contratados. Esta língua de contacto tornou-se língua materna das gerações seguintes, já nascidas em São Tomé, e caracterizou-se como um português reestruturado com traços das línguas maternas africanas e do Português de São Tomé (PST) (Hagemeijer 2009a). Atente-se para os exemplos 1) e 2):

1) Elé ni Angola é febere, quando chegô aqui élé pariu ami, cabô, nõ pariu mase, mase.
(Ele (,) em Angola era febre, quando chegou aqui ela pariu-me. Depois não pariu mais.)

(BAXTER, 2002)

2) Nós é tonga de Bairro da Liberdade.
(Nós somos Tongas do Bairro da Liberdade)

(Dados orais, TINY COSME, 2021)

O Português dos Tongas não se trata, todavia, de uma língua homogênea uma vez que esteve confinada ao contexto específico das sanzalas das roças de Água Izé, Monte Café, e Agostinho Neto (ROUGÉ, 1992). A maior liberdade de movimentação e fixação das populações ao longo de todo o território, bem como o maior acesso à educação, conduzirá a uma possível dissolução desta variedade do português.

3. Português em São Tomé e Príncipe

Após a proclamação da independência do país, em 12 de julho de 1975, o Português ganhou o estatuto de língua oficial, conservando, até então, a exclusividade do estatuto. O crescimento do número de falantes do português, por sua vez, tem sido diretamente proporcional ao crescimento da população ao longo da história do país.

QUADRO 1
Distribuição da população por língua

	População	Português	Forro	Lung'le	Angolar	Kabuverdian	Outras
1981	96.661	60.519	54.387	1.533	-	-	8.180
1991	117.504	94.907	69.899	1.558	-	-	12.781
2001	140.365	136.085	88.621	3.302	-	-	17.612
2012	178.739	175.879	64.794	1.787	11.797	15.193	4.224

Fonte: Adaptação dos dados do INE (2022)

De acordo com os dados do Quadro 1, verifica-se que nos trinta anos após a independência, a população do país praticamente duplicou, registrando, a mesma fonte o número de 197.700 habitantes no ano de 2017 (INE, 2022). No que concerne às línguas crioulas, o Forro e o Lung'le registam uma preocupante queda do número de falantes, tornando-se ainda mais preocupante quando se considera o curto período de tempo em que esta se assinala. Quanto ao Angolar e ao Kabuverdiano, foram considerados apenas no censo de 2012, pelo que não é possível um olhar comparativo, porém, acredita-se que o número de falantes estará integrado na opção de *outras línguas*.

O Português, contrariamente às demais línguas, tem sofrido um aumento exponencial do número de falantes, encontrando-se, atualmente, na ordem dos 98% a população que afirma ser falante desta. É, portanto, possível concluir que

"... em 1981, havia mais monolíngues em português e em Forro, ao passo que o censo mais recente reforça a tendência para um crescente monolinguismo em português. Por outras palavras, a situação de diglossia relativamente estável que se verifica no tempo colonial, em que o português era a língua dos domínios altos e o forro a língua dos domínios baixos, deu origem a um contexto de diglossia instável que se caracteriza pela generalização do português em todos os domínios, em detrimento dos crioulos." (GONÇALVES; HAGEMEIJER, 2015)

Relaciona-se esta tendência para a generalização do português em detrimento dos crioulos, mormente em relação ao Forro, com a atribuição, no tempo colonial, de estatuto de língua inferior associada aos descendentes dos escravizados. Foi, por isso, e de modo pejorativo, apelidado de dialeto, termo depreciativo, eivado de preconceitos, que, de forma ignorante e por ignorância, continua, ainda e também, a ser aplicado por muitos falantes, de forma transversal às diferentes classes sociais.

O período pós-independência não trouxe consigo mudanças significativas e estruturantes, expectáveis num contexto de afirmação identitária, em que as línguas se posicionam enquanto fator identitário e de unidade nacional. As tentativas de valorização e aproximação às línguas numa perspetiva progressista de construção de uma identidade nacional revelaram-se pouco objetivas e menos eficientes e eficazes.

“Com efeito, no período imediato à independência, essa mesma elite foi promotora do sentido, muito monolítico, de patriotismo tendo concentrado na sua língua materna (a portuguesa) a recusa do que considerava serem valores do colonialismo, interessada como estava numa aproximação às camadas mais desfavorecidas, dentro de um programa de valorização das culturas nacionais, de uma política de pseudo-igualdade e na miragem de uma “sociedade sem classes”. Iniciativas mal conduzidas, derivadas de uma grande dose de voluntarismo populista, embora compreensíveis no contexto.” (MATA, 2009)

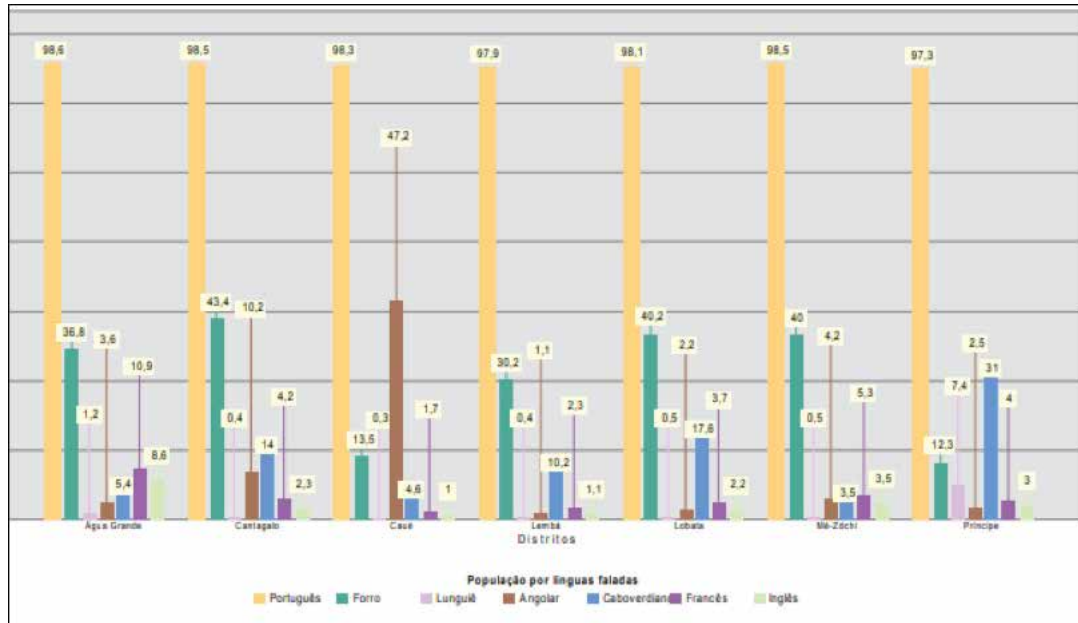
A hegemonia funcional resultante da atribuição do estatuto único de língua oficial conduziu à sua manutenção enquanto língua de poder e de acesso ao poder, língua materna da elite, tendo-se proliferado por meio da massificação do ensino e da alfabetização (MATA, 2009).

4. O Português de São Tomé e Príncipe

4.1. Contextualização

O Português, outrora caracterizado como de domínio exclusivo do colonizador, é hoje parte integrante do património linguístico nacional coletivo. A massificação do ensino, associada à maior liberdade de movimento e fixação por parte da população, bem como o êxodo das roças para as zonas urbanas e periurbanas conduziram ao maior acesso à língua justificando o alto índice percentual de falantes.

TABELA 2
Distribuição (%) da população por línguas faladas segundo distrito (2012)



Fonte: INE (2022)

A distribuição patente no Quadro 2 aponta para o predomínio do Português, totalizando uma média de cobertura de aproximadamente 98% em todos os distritos do país, bem como na Região Autónoma do Príncipe. O distrito de Água Grande, o maior do país em termos de densidade populacional, correspondendo à localização mais urbana, apresenta, em termos comparativos, baixa frequência de falantes das línguas crioulas, onde, em exclusão do Forro, não apresentam percentagem superior a 10% de falantes.

4.2. A norma-padrão e o Português popular

A hegemonia do Português quanto ao contexto linguístico são-tomense é consensual e transversal aos diversos autores, sendo este considerado o único país africano de língua oficial portuguesa (PALOP) onde a maioria da população tem o Português como primeira língua.

Impõe-se, todavia, questionar de que Português se fala quando se afirma a hegemonia deste. São frequentes as afirmações de que existem apenas duas variedades do Português, a europeia e a brasileira; do mesmo modo que se propala, igualmente e reiteradamente, que a norma-padrão existente em STP é a do português europeu (PE).

A norma-padrão é, naturalmente, a selecionada para ser ensinada no âmbito da alfabetização das crianças, é a transmitida em contexto escolar, onde o acesso aos professores é direto e reiterado, é a utilizada em contexto académico, entre os docentes e discentes. Considera-se que esta norma é, similarmente, veiculada nos meios de comunicação social, oral e escrito, sendo a rádio o meio privilegiado para fazer chegar todo o tipo de informação diária à população, com maior cobertura nacional e de mais fácil acesso, pois é o que menor

custo acarreta para o cidadão, no caso específico de STP. Os documentos de carácter oficial, nomeadamente documentos ministeriais, projetos e resoluções da Assembleia Nacional, informações constante no Diário da República, imprensa escrita, livros escolares e de escritores e autores nacionais são apontados como difusores do Português normativo. Atente-se para o conjunto de frases abaixo, resultado de recolhas orais e escritas, em contexto de imersão e exposição diária ao Português falado em STP, maioritariamente em Água Grande e Mé Zochi, os dois distritos mais populosos do país, entre os anos de 2019 e 2022.

1) “Norma Europeia”

- a) Toda a gente não pode fazer (construir) pontes. (PSTP¹)
(Nem toda a gente pode construir pontes (PE²))
- b) Todos não são corruptos. (PSTP)
(Nem todos são corruptos. (PE))
- c) O lápis da fulana perdeu. (PSTP)
(A fulana perdeu o lápis. (PE))
- d) A caneta perdeu. (PSTP)
Kanêta plêdê. (Forro)
(Caneta perdeu (tradução literal))
(Alguém perdeu a caneta (PE))
- e) O escritor enganou. (PSTP)
(O escritor enganou-se (PE))
- f) Isto é uma falta de respeito e carregada de **wêchaismo**³. (PSTP)
- g) Cataplana, o a[r]oz bom.
- h) Para (verbo) / para (preposição) – Homonímia
- i) Deus nos abençoe à todos.

Os exemplos acima introduzidos são enunciados produzidos nos meios onde se veicula o uso da norma-padrão, mormente a norma europeia. Foram produzidos por locutores de rádio e pivôs de telejornais, durante debates parlamentares, por alunos e professores em sala de aula e conferências e nos restaurante e cafés frequentados por parte da elite letrada.

É possível verificar um conjunto de singularidades, partilhadas por um vasto número de falantes. No par de frases a) e b), o advérbio de negação **não** ocorre após o quantificador universal **todos** em contexto onde a norma europeia requereria a ocorrência do constituinte

¹ PSTP - Português de São Tomé e Príncipe

² PE – Português Europeu

³ Wêchaismo – Inveja

de negação de universalidade **nem**, em posição inicial na oração.

Os enunciados c) e d) refletem a influência do Forro no Português onde o falante faz a transferência da sintaxe do Forro, numa tradução direta de “Kanêta plêdê”.

O substantivo **wêchaismo** deriva do adjetivo **wê xa**⁴ do Forro. Esta entrada está já registada no Vocabulário Ortográfico São-tomense da Língua Portuguesa⁵ (VOSLP), enquanto vocabulário específico de STP.

Por fim, a frase h) resulta do contexto de sala de aula onde foi solicitado exemplos de palavras homónimas, a alunos de ensino superior. O par opositivo indicado: **para** (verbo) vs **para** (preposição), atestado por todos os alunos presentes na sala, demonstra que a nível fonético-fonológico há, também, particularidades do PSTP, divergentes da norma do PE, pois a vogal /a/ quer em posição tónica, quer na átona, é produzida por estes como vogal central baixa - [a]. No PE padrão esperar-se-ia a ocorrência de oposição entre a vogal central baixa [a] e a vogal central média [e], na posição tónica.

2) Português Popular (PP)

a) Domingo eu molhei chuva! (PP)

Ja jingo n monha ãa suba!

(Domingo eu molhei-me na chuva (PE))

b) Eu piquei pinha. (PP)

Fui picado por uma pinha. (PE)

c) Eu próprio votei afabal... votei afabal. (PP)

d) Bom vinho, mau dele. (PP)

e) Tenho que estar com muita postura, tenho que estar posturada no trabalho.
(PP)

As frases em 2) são o resultado da recolha de produção oral, em contexto não formal, menos elitizado, onde se veicula o que chamaremos de Português popular⁶.

A estratégia de tradução direta do Forro para o Português é também produtiva no PP, observável na frase a), onde a influência é visível na transferência da estrutura sintática, resultando na omissão do pronome reflexivo **me**.

A palavra **afabal** em c) é vocabulário específico do PSTP, já integrado no VOSTLP, que na frase acima tem o valor de adjetivo, significando “de forma gratuita/sem cobrar contrapartida”.

Na frase e) é possível verificar a criatividade do falante, que no uso de estratégias possíveis na língua, partindo do nome **postura** construiu o verbo e a correspondente forma participial **posturada**.

⁴ Wê xa - Invejoso

⁵ Vocabulário Ortográfico São-tomense da Língua Portuguesa – Instrumento já apresentado e homologado nacionalmente. Aguarda introdução na plataforma do Vocabulário Ortográfico Comum da Língua Portuguesa - CPLP

⁶ Autores como Fernanda Pontífice e Albertino Bragança denominam de Português oral

5. Considerações finais

A norma apontada como padrão no contexto de STP é produzida i. pela elite escolarizada na época colonial e seus descendentes, ii. pelos falantes formados e/ou com vivência de longos períodos no exterior do país (nomeadamente em Portugal), iii. pelos cidadãos mais escolarizados (ao nível do ensino superior), tendo feito o percurso internamente e cuja vivência permite o contacto constante com contextos privilegiados. Nesse sentido, pode-se afirmar que a percentagem de falantes detentores do Português padrão é, necessariamente, reduzido, encontrando-se circunscrito a grupos perfeitamente identificados. Todavia, com base nos dados apresentados, referentes à produção oral de falantes do PSTP padrão, verifica-se que os enunciados não são conducentes com o PE, distinguindo-se a diferentes níveis, designadamente no que concerne a aspetos sintáticos, lexicais e fonético-fonológicos. É falacioso, portanto, caracterizar este grupo como falantes do PE, tendo este que ser enquadrados, inevitavelmente, numa outra realidade.

Outrossim, os falantes do PP constituem a parte massiva da população, não integrados nos três grupos acima identificados. O PP é, de imediato, relacionado exclusivamente às camadas menos privilegiadas e menos escolarizadas, procurando-se estabelecer uma relação direta entre o PP e a classe social. Não obstante, quando se emerge na realidade, as evidências, por vezes, contrariam as pré-concepções. É prova desta afirmação a situação linguística dos alunos da escola portuguesa em São Tomé e Príncipe, uma escola privada, onde concentra os filhos dos grupos falantes da norma padrão, descritos nos pontos i., ii. e iii. Seria de esperar que entre estes a língua em uso nos grupos, nos intervalos, fora do contexto de sala de aula, fosse a norma padrão, quer pela influência do ambiente familiar, quer pelo *input* em sala de aula, uma vez que é ensinada a norma europeia, por uma maioria de professores portugueses. Contudo, não é o que acontece. É, antes, notório um uso da língua mais próximo do PP. Este contexto de contacto linguístico inter-geracional constitui uma base interessante e útil na perspetiva do estudo do PSTP.

A escolarização consubstancia a estratégia mais eficiente para o ensino e aprendizagem das normas instituídas. Mais uma vez, neste particular, é sistematicamente propalado que a norma ensinada nas escolas de STP corresponde ao PE. No entanto, é amplamente conhecida a complexidade da problemática do sistema de ensino, sobretudo no que se refere a competências específicas, correspondentes aos diferentes níveis de aprendizagem. Ao nível do ensino da Língua Portuguesa nas escolas, as complexidades não são menores. Para além do défice de preparação pedagógica dos professores, são, também, evidentes as fragilidades técnicas refletidas na qualidade do ensino, tendo-se criado, ao longo dos tempos, um circuito de complexidade linguística. Os alunos que ingressam nas escolas, desde o ensino pré-primário, têm o PP como L1. Esta é a língua que a criança adquire no ambiente familiar e é esta que circula no seu contexto sociocultural. Quando estes atingem a fase da escolarização, encontram professores, eles próprios falantes PP, com maior ou menor intensidade. No entanto, são encarregues de ensinar a norma do PE, que os próprios não dominam. O resultado é uma realidade linguística própria e um ambiente com as condições propícias que contribuem para a fase inicial da consolidação de uma variedade do PSTP, cuja norma padrão deverá, necessariamente, refletir estas idiosincrasias.

Nesse português insularmente re-ontologizado vem-se alojando a matriz cultural de uma variante cuja norma, embora não definida e até sistematizada, se percebe já como uma realidade perfeitamente identificável como marcas específicas na expressão. (MATA, 2009)

Em tempos de afirmação do pluricentrismo da Língua Portuguesa, urge que se façam estudos consentâneos e análises críticas à realidade linguística efetiva do país, para que se promovam políticas direcionadas de reconhecimento e institucionalização da variedade do português de STP e da sua respetiva norma-padrão, numa perspetiva congregadora, despida de preconceitos e de amarras de olhares neocoloniais.

6. Referências bibliográficas

Bakker, Peter; Daval-Markussen, Aymeric; Parkvall, Mikael; Plag, Ingo. 2011. *Creoles are typologically distinct from non-creoles*. Journal of Pidgin and Creole Languages 26(1): 5-42.

Caldeira, Arlindo. 2004. *Rebelião e outras formas de resistência à escravatura nas ilhas do Golfo da Guiné* (séculos XVI-XVIII). Studia Africana 7. 101-136.

Caldeira, Arlindo. 2006. *Uma ilha quase desconhecida. Notas para a história de Ano Bom*. Studia Africana – Revista Interuniversitária d'Estudis Africans 17. 99-109.

Cosme, Abigail Tiny. 2014. *As Relações Filogenéticas entre os Crioulos do Golfo da Guiné*. Dissertação de Mestrado (em Linguística), Faculdade de Letras da Universidade de Lisboa. Disponível em https://repositorio.ul.pt/bitstream/10451/18351/1/ulfl175416_tm.pdf. Acesso em 10 de outubro de 2022

Ferraz, Luiz Ivens. 1974. *A linguistic appraisal of Angolar*. In Memoriam Antonio Jorge Dias, Vol. 2, 177-186. Lisbon: Instituto de Alta Cultura/Junta de Investigações Científicas do Ultramar.

Ferraz, Luiz Ivens. 1976. *The origin and development of four creoles in the Gulf of Guinea*. African Studies 35(1). 33-38.

Ferraz, L. 1979. *The creole of São Tomé*. Johannesburg: Witwatersrand University Press.

Ferraz, Luiz Ivens. 1987a. *The liquid in the Gulf of Guinea Creoles*. African Studies 46(2): 287-95.

Ferraz, L. 1987b. *Portuguese creoles of West Africa and Asia*. In Gilbert, Glen (ed.), *Pidgins and creole languages: Essays in memory of John E. Reinecke*, 337-60. Honolulu: University of Hawaii Press.

Gonçalves, R. & Hagemeijer, T. 2015. *O português num context multilingue: o caso de São Tomé e Príncipe*. Revista Científica da Universidade Eduardo Mondlane, Série: Letras e Ciências Sociais.

Günther, Wilfried. 1973. *Das Portugiesische Kreolisch der ilha do Príncipe*. Marburg an der Lahn: Marburger Studien zur Afrika- und Asienkunde.

Hagemeijer, Tjerk. 2009. *As línguas de S. Tomé e Príncipe*. Revista de Crioulos de Base Lexical Portuguesa e Espanhola, 1:1, 1-27.

Hagemeijer, Tjerk. 2011. *The Gulf of Guinea creoles: genetic and typological relations*. Journal of Pidgin and Creole Languages. 26:1, 111-154.

Lorenzino, Gerardo. 1998. *The Angolar Creole Portuguese of São Tomé: Its grammar and sociolinguistic history*. Ph.D. dissertation, City University of New York.

Mata, Inocência. 2009. *Latitudes – Cahiers Lusophones (Paris). Dossier spécial: São Tomé e Príncipe: culture et société*. N°36, 3-11.

Maurer, Philippe. 1995. *L'Angolar: Un créole afro-portugais parlé à São Tomé*. Hamburg: Helmut Buske Verlag.

Maurer, Philippe 1999. *El verbo locativo poner en santomense, principense y angolar*. In Klaus Zimmermann (ed.), *Lenguas criollas de base lexical española y portuguesa*, 89-100. Frankfurt/Madrid: Vervuert/ Iberoamericana.

Maurer, Philippe. 2009. *Principense — Grammar, texts, and vocabulary of the AfroPortuguese creole of the island of Príncipe*. London: Battlebridge Publications.

McMahon, April & Robert McMahon. 2003. *Finding families: Quantitative methods in language classification*. Transactions of the Philological Society 101. 7–55.

Michaelis, Susanne; Maurer, Philippe; Haspelmath, Martin; Huber, Magnus (eds.). 2013. *Atlas of pidgin and creole language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Michaelis, Susanne; Maurer, Philippe; Haspelmath, Martin; Huber, Magnus (eds.). 2013. *The survey of pidgin and creole languages*. Volume 2: Portuguese-based, Spanish-based, and French-based Languages. Oxford: Oxford University Press.

Negreiros, Almada José. 1895. *Historia Ethnographica da ilha de S. Tomé*. Lisboa.

Schang, Emmanuel. 2000. *L'émergence des créoles portugais du Golfe de Guinée*. Tese de Doutorado, Universidade de Nancy.

Seibert, Gerhard. 2007. *Angolares of São Tomé island*. In Philip Havik & Malyn Newitt (eds.), *Creole societies in the Portuguese colonial empire*, 105-126. Bristol: Bristol University Press.

Sousa, Celso Batista de. 1990. S. Tomé e Príncipe. *Do descobrimento aos meados do século XVI*. Dissertação de Mestrado, Universidade de Lisboa.

Instituto Nacional de Estatística (ine.st). Acesso em 10 de outubro de 2022

A língua portuguesa no espaço dos Países Africanos de Língua Oficial – Uma reflexão sobre Cabo Verde

Adelaide Monteiro

Sobre o autor

Instituto do Património Cultural / Ministério da Cultura e das Indústrias Criativas; ponto focal da Comissão Nacional de Cabo Verde – IILP, linguista pela Faculdade de Letras, Mestre em letras neolatinas pela Faculdade de letras da Universidade Federal do Rio de Janeiro, doutoranda em linguística na Faculdade de Letras da Universidade de Lisboa
adelaide.t.monteiro@gmail.com

Resumo

Este artigo responde ao convite para participação na conferência A língua portuguesa no espaço Países Africanos de Língua Oficial Portuguesa - PALOP - Nesta apresentação, tendo em conta o público alvo, maioritariamente estudantes universitários de letras da Universidade de Brasília, descrevemos o uso da língua portuguesa em Cabo Verde, arquipélago de 10 ilhas, situado na costa Ocidental Africana que a par do seu povoamento e colonização portuguesa, enfrentou também o aparecimento de uma língua nativa, de nome língua cabo-verdiana, mas que o contexto histórico da sua formação permite que se lhe chame crioulo, que é usado ao lado da língua portuguesa, desde os meados do Séc. XVI. Duas línguas que partilham um grande número de léxico comum, mas que se afastam nas suas estruturas morfologias e sintáticas. A prática linguística do cabo-verdiano no seu dia-a-dia é marcada pelo das duas línguas, cuja escolha é condicionada pelo contexto formal e informal e pelas modalidades oral versus escrita, que o obriga a usar a sua língua materna, o cabo-verdiano, língua não oficial ou a sua língua oficial, a língua portuguesa. Faremos uma abordagem de gestão de línguas em situação de contactos linguístico, entre a língua cabo-verdiana e a língua portuguesa, que pelo estatuto desempenham funções diferentes, mas a prática linguística impulsiona a definição de uma outra política de língua e uma gestão das duas línguas presentes no país mais próxima da realidade sociolinguística de Cabo Verde. Para falar de política de língua recorreremos ao sociolinguista Jean-Louis Calvet e Henri Boyer, quadro teórico que sustentou também a nossa dissertação de mestrado. A nossa análise apoia-se nas práticas linguísticas dos falantes em domínios informais e formais, onde nós utilizamos uma metodologia de observação não participante e ainda, nos apoiamos numa bibliografia sobre a Cultura e Educação em Cabo Verde, bem como as legislações sobre a educação e língua.

Palavras-chaves

Cabo Verde, língua portuguesa, língua crioula cabo-verdiana.

1. Introdução

Cabo Verde é um arquipélago de 10 ilhas, na costa Ocidental Africana e é um dos países que tem a língua portuguesa como sua oficial. Muitos dos outros países de língua oficial portuguesa já tinha uma língua, ou línguas a quando da chegada dos colonizadores portugueses. A par do seu povoamento e colonização portuguesa, enfrentou também o aparecimento de uma língua nativa, a língua cabo-verdiana¹. Para falar da vida da língua portuguesa em Cabo Verde há que falar da língua cabo-verdiana, pois a língua enquanto produto social ela não existe sem os seus usuários e no caso, os usuários da língua portuguesa, que é a língua oficial e de ensino, eles usam a sua língua materna, que é a língua crioula cabo-verdiana. A prática linguística do cabo-verdiano no seu dia-a-dia, é condicionada pelo contexto formal e informal e pelas modalidades oral versus escrita, que o obriga a usar a sua língua materna o cabo-verdiano, língua não oficial e a sua língua oficial, a língua portuguesa. Faremos uma abordagem de gestão de línguas em situação de contacto linguístico, entre a língua cabo-verdiana e a língua portuguesa, que pelo estatuto desempenham funções diferentes, mas a prática linguística indica uma possível mudança na atual política de língua e uma gestão das duas línguas presentes no país mais próxima da realidade sociolinguística de Cabo Verde.

Nas situações de comunicação oral foram observadas as salas de aula, mesas redondas, palestras, debates parlamentares, o dia-a-dia nos serviços administrativos, a comunicação social nacional (Rádio e Televisão) e a comunicação em ambientes familiares e na rua. Quanto às situações de comunicação escrita, foram consultados jornais impressos e eletrónicos, atas da Assembleia Nacional de Cabo Verde, ofícios administrativos, literatura disponível, anúncios escritos na televisão e *emails*; situações de comunicação oral são as observações que fizemos de situações de diálogos em ambientes familiares, na rua e no Parlamento. Consideramos situações de comunicação em presença, as em que os interlocutores estão face a face e à distância os feitos através de *email*, pela Rádio e pela televisão.

2. Desenvolvimento

Vamos, a seguir, ilustrar os contextos do uso das duas línguas nas modalidades oral e escrita, em diferentes contextos situacionais e comunicacionais.

¹ Designada comumente pelos próprios falantes como crioulo, tendo em conta o contexto histórico do aparecimento das línguas designadas crioulas. Com a divulgação dos estudos sobre esta língua e o conhecimento de que a expressão crioula relativamente à expressão *língua* pode ser interpretado como pejorativo, tem-se por um lado optado por *língua cabo-verdiana*, ou *língua crioula cabo-verdiana*.

2.1. Contexto comunicacional e situacional do uso das duas línguas

2.1.1. Contexto de comunicação oral

2.1.1.1. Família

Durante a observação que fizemos, notamos o funcionamento das duas línguas em Cabo Verde, verificámos que estando em casa, em um ambiente familiar, o locutor cabo-verdiano, escolhe na maior parte das vezes a língua cabo-verdiana. Encontramos poucas situações em que a língua portuguesa era a língua materna.

No entanto, registamos que as histórias tradicionais, em qualquer ambiente são em cabo-verdiano bem como anedotas, adivinhas e outros conversas de convívios informais.

2.1.1.2. Escola (Públicas)

A grande maioria dos professores fala cabo-verdiano entre si não só na sala do professor, mas também fora dela. Durante as reuniões de coordenação, normalmente os coordenadores iniciam a reunião em português, mas momentos depois continua prevalecendo o diálogo em língua cabo-verdiana. Notamos ainda que os professores na faixa etária mais avançadas, anos são os que mais fazem uso da língua portuguesa quer com os colegas quer com os alunos e também nas reuniões. Os conteúdos programáticos são transmitidos em língua portuguesa. É o próprio relatório da UNICEF sobre as dificuldades e possibilidades dos professores e dos alunos no processo de Ensino-Aprendizagem citado em (2003) no Plano Estratégico para a Educação, M.E.V.R.H.² (p.40) que diz que "O ensino dos conteúdos, na maioria das aulas observadas, é marcado quase sempre por um ritual em que o silêncio, a submissão, a sacralização do saber e do professor, são os seus elementos constitutivos."

2.1.1.3. Convívio social e eventos culturais

Como nos outros sectores, é possível encontrar a utilização das duas línguas. Os grupos teatrais nacionais apresentam as peças em português ou em cabo-verdiano e como nos explicou um dos diretores de um grupo teatral, muitas vezes depende do tema a abordar e são os próprios autores que acabam por decidir.

2.1.1.4. Atividades religiosas

Nos últimos anos tornou-se frequente as celebrações religiosas serem proferidas em língua cabo-verdiana o que há duas décadas atrás não era possível.

² Ministério da Educação e Valorização dos Recursos Humanos

2.1.1.5. Nos Serviços Administrativos

As duas línguas aparecem com muita frequência e dependo dos interlocutores pode-se ouvir o cabo-verdiano ou o português. Notou-se mais tendência do uso do português entre a classe dirigente, mas atualmente, os quadros mais jovens, mesmo estando em posição de chefia recorrem muitas vezes ao cabo-verdiano, durante as reuniões.

2.1.1.6. Comunicação social

Na Rádio e na Televisão Nacional de Cabo Verde a apresentação das notícias é feita em português. No entanto, nas reportagens, entrevistas, programas culturais e debates é frequente ouvirmos o jornalista a se dirigir aos entrevistados em língua cabo-verdiana bem como a tendência de ser em português quando se trata de algum intelectual ou político. Muitas vezes, em anúncios ou avisos, o texto passa na tela da televisão em português, mas a voz *off*, é em cabo-verdiano.

2.1.1.7. Publicidade

Este é um setor cujos textos eram sobretudo em português. Ultimamente, é frequente ouvir na Rádio e na Televisão públicas ou privadas textos em língua portuguesa e em língua cabo-verdiana. A própria empresa de comunicações telefónicas CABO VERDE TELECOM, S.A tem apresentado planos para telemóveis cujos nomes são em cabo-verdiano. Empresas nacionais e estrangeiras, apresentam publicidades com textos escritos em Português ou em cabo-verdiano.

A divulgação de atividades culturais e desportivas, bem como políticas são feitas em cabo-verdiano e Português. Mas o serviço oficial da Rádio Nacional que divulga informações de emprego, comunicações de falecimento, é feito em português. No entanto, no mesmo programa pode-se ouvir convites em crioulo, para bailes de grupos musicais ou atividades teatrais.

2.1.2. Contexto situacional de comunicação escrita

2.1.2.1. Textos oficiais

Como dissemos no início, na modalidade escrita predomina o uso da língua portuguesa. Não obstante, começam a aparecer mais textos literários escritos em língua cabo-verdiana. Em textos formais, como as atas das sessões plenárias, datadas a partir de 1999. Nas revistas podemos encontrar, esporadicamente, títulos em cabo-verdiano, sobretudo quando se trata de matérias de festas tradicionais. Nos géneros jornalísticos, entrevistas e reportagens, encontramos textos dos entrevistados que foram traduzidos da língua cabo-verdiana para o

português. No entanto, é comum encontrar várias expressões do quotidiano cabo-verdiano em qualquer jornal ou revista nacional quer em versão eletrónica quer impressas.

2.1.2.2. No sistema educativo

O português foi sempre a língua oficial do sistema educativo e por isso, a única utilizada na modalidade escrita, enquanto que o cabo-verdiano nunca foi utilizado no sistema de ensino na modalidade escrita a ponto de os alunos desconhecerem a existência do Alfabeto Unificado para a Escrita do Cabo-verdiano, aprovado em pelo governo em 2009. Não há alfabetização em língua cabo-verdiana.

2.1.2.3. Na literatura

Na área da literatura, as publicações são maioritariamente em língua portuguesa. No entanto, aparecem sempre expressões em cabo-verdiano nos textos, sobretudo na boca das personagens, assim, como é a língua cabo-verdiana que se usa no caso de sessões de contação de histórias, que eram contadas às crianças no final do dia. Existe, hoje, versões bilingues, de algumas destas histórias³. Tem aparecido ultimamente muitas edições de poesia e romance em língua cabo-verdiana. Existem registos escritos de poesias em cabo-verdiano, desde finais de séc- XIX. Muitos poemas dos escritores cabo-verdianos são hoje letras de músicas que aparecem escritas em cabo-verdiano nos discos e cd's. Aliás, os discos e cd's são os suportes que mais têm divulgado o cabo-verdiano na modalidade da escrita. Devido a ausência de uma normalização da escrita da língua cabo-verdiana, pode-se verificar a alternância no uso das grafias. Se a grande maioria dos textos literários, nos géneros de romance, poesia e são escritos em língua portuguesa o mesmo já não se pode dizer das letras das músicas, inclusive música religiosas.

2.1.2.4. Nos topónimos e onomástica

Encontramos inscrições em português, mas os falantes sempre as pronunciam em cabo-verdiano. O nome das ruas aparece sempre em português. Em muitos casos, o próprio nome do indivíduo tem versão bilingue. Por exemplo, "João" é frequente que alguém o chame de "Djon". No entanto, o que aparece na Certidão Narrativa de Nascimento é o nome em português ou em outra língua, mas nunca em cabo-verdiano.

2.1.2.5. Comunicação nas redes sociais e chats

Na comunicação instantânea via Internet, usa-se cada vez a língua cabo-verdiana. Os *emails* que inicialmente sempre apareciam em português começam a surgir, em cabo-

³ Nomeadamente, as histórias do "Ti Lobo e Chibinho"

verdiano. A camada mais jovem dificilmente recorre ao português para escrever *email* aos amigos. Facto interessante anotado aqui é a convivência, na modalidade escrita, das variantes linguísticas existentes na própria língua cabo-verdiana. Perante a ausência de uma norma ortográfica bem como de uma variedade padrão, os internautas escrevem de acordo com a variedade que eles falam.

Esta descrição mostrou-nos que, neste momento, se assiste a uma redistribuição do uso das duas línguas. Já não se pode falar em exclusividade do uso da língua portuguesa na modalidade escrita e, em espaços oficiais. Reconhecemos o uso exclusivo da língua portuguesa em documentos oficiais⁴.

2.1.3. A língua utilizada em situações formais de comunicação

Outra constatação são as características gramaticais da modalidade oral da língua utilizada na comunicação social oral e em espaços oficiais, como o Parlamento, nas reuniões de trabalho ou encontros cujos interlocutores ocupam altos cargos na Administração pública, nota-se um predomínio do uso do português. No entanto, uma observação atenta mostra a existência de grandes interferências do cabo-verdiano no português, interferências duramente criticadas como sendo erros gramaticais no português.

Até a independência de Cabo Verde, a política linguista vigente era de repressão do uso da língua cabo-verdiana em espaços formais, com destaque para o ensino e o uso obrigatório do Português, inclusive durante o intervalo, no recinto escolar. (" A nível da instrução primária, a primeira medida data de 1849, tendo constado do " Programma Interino da Escola Principal de Instrução Primaria" que " na Escola só é permitido falar portuguez, o dialecto crioulo é absolutamente prohibido." (Brito-Semedo: 2006:73) O autor ainda diz-nos que "Posteriormente, em 1920, o Governador Almeida Maia Magalhães (1919-1921), tendo notado, na sua visita às escolas das diferentes ilhas, que nalgumas delas o ensino se fazia ainda em crioulo, ou misto de crioulo e português, e sido informado pelos inspector escolar que "esse abuso" tinha sido levado ao ponto de alguns interrogatórios dos exames se realizarem nessa língua, houve por bem decidir " proibir expressamente o uso do crioulo nas escolas e determinar que a inobservância desta ordem seja considerada desobediência e, como tal, punidos todos os professores que infringirem ". Também no ensino secundário a política linguística em causa, era de opressão quanto ao uso da língua cabo-verdiana. Brito-semedo (2006:73) cita o Regulamento Interno do Liceu Central D. Henrique, em S. Vicente, referente ao ano de 1932, que estabelecia no artigo 3º o seguinte: "É expressamente proibido falar crioulo dentro do edifício do Liceu."

Após esta explanação do uso das línguas portuguesa e cabo-verdiana, em Cabo Verde pode-se concluir que a língua portuguesa em Cabo Verde é usada predominantemente na modalidade escrita e que estamos perante uma sociedade com grandes potencialidades de construção de um bilinguismo. No entanto, a política linguística vigente não tem criado condições para tal.

Considerando que a **política linguística** é toda a escolha consciente efectuada no domínio das relações entre a língua e a vida social e mais particularmente entre a língua e vida

⁴ Com exceção das atas da Assembleia Nacional de Cabo Verde.

nacional enquanto que a **planificação** linguística é a implementação dos meios necessários para a aplicação de uma política linguística., no conceito de Calvet (1987:154-155) e, ainda, de acordo com Boyer (2001:70) que diz que a gestão oficial de situações de monolinguismo o ou plurilinguismo não é apenas uma gestão, puramente, linguística. Ela depende de situações históricas, sócio-étnicas, económicas e demográficas. Ele define a **política linguística** como toda a acção de um Estado que designa escolhas, orientações e objectivos deste Estado em relação à gestão das línguas quer em situações de plurilinguismo quer em situações de unilinguismo.

Se por um lado fazemos uma leitura de uma política linguística favorável à língua portuguesa, criando um modelo monolingue, por outro tomamos boa nota a disseminação do uso da língua cabo-verdiana para além do espaço familiar e informal e da oralidade. Mas também é verdade e somos obrigados a admitir que perante esta manifestação do uso da língua materna, há também uma auto-repressão da parte de muitos cabo-verdianos em usar estas duas línguas, a saber, o uso da língua portuguesa em espaços informais e na oralidade e no uso da língua cabo-verdiana em espaços formais e na modalidade escrita. Esta representação⁵ que os falantes têm das suas práticas, repercutem na própria tomada de decisão sobre a política e no intimidam o uso dessiminado seja da língua portuguesa seja da língua cabo-verdiana.

Calvet (1999) propõe o cruzamento das práticas linguísticas com as representações dos locutores o que acaba por ser o cruzamento de o que os locutores pensam da sua maneira de falar (segurança / insegurança formal) e o valor que eles dão ao que eles falam (segurança estatutária). Cruzam-se apenas os dados provenientes dos locutores, que ele questiona como eles se auto-avaliam.

Reconhece-se o uso da língua portuguesa, língua de ensino, em espaços formais e por pessoas cultas e supostamente detentoras de status social versus o uso da língua crioula, proibido no recinto escolar, permitida em casa e em espaços informais, fez criar e cristalizar um imaginário linguístico pejorativo para a língua cabo-verdiana que ainda hoje persiste em algumas cabeças ainda que as práticas linguísticas demostrem o contrário. os discursos epilinguísticos⁶ feitos durante a colonização e ainda hoje reproduzidos por alguns professores ou pessoas com reconhecimento social de que é a língua portuguesa que dá acesso ao conhecimento e sucesso profissional contribuem para criar no falante situações de insegurança linguística no uso das duas línguas e uma trofia no desenvolvimento de cada uma delas. Mas se levarmos em conta o modelo gravitacional de Calvet, que diz que cada falante deve dominar pelos menostres línguas (a sua língua nacional, a do Estado e uma Interncional) para ser um cidadão lingusticamente competente não temos dúvidas de que a língua portuguesa ganhará mais usuários cabo-verdianos. Uma outra consequência será o desenvolvimento e asunção de uma variedade cabo-verdiana de língua portuguesa cuja investigação começa a despontar com os projetos da Cátedra eugénio Tavares da Língua Portuguesa na Universidade de Cabo Verde e do Centro de Linguística da Universidade de Lisboa.

⁵ Entenda-se por **representações** o postulado por Calvet (1999): "a maneira como os locutores pensam as suas **práticas linguísticas** sabendo que por práticas linguísticas entende o mesmo autor que que são o que os locutores produzem; a maneira como eles falam mas também a maneira como eles adaptam as suas práticas às diferentes situações de comunicação, como por exemplo, adotar as práticas às expectativas dos interlocutores.

⁶ Calvet (19999) compreende, que os discursos epilinguísticos são significados do imaginário linguístico e este faz parte (ou é equivalente) das atitudes.

3. Conclusão

Como isso será possível? Há que ultrapassar os discursos epilinguísticos, e desenhar uma política linguística favorável ao desenvolvimento das línguas presentes no País. Tecnicamente, a língua portuguesa pode ganhar 1 milhão de falantes cabo-verdianos diretos e muitos outros indiretamente, mas dificilmente os ganhará se não houver uma política linguística educativa que não deixe de forma a língua materna dos cabo-verdianos.

4. Referências Bibliográficas

Brito-Semedo, Manuel (2006) – *A construção da identidade nacional* :análise da imprensa entre 1877 e 1975. INBL

CALVET, Louis-Jean. *La guerre des Langues et les politiques lingusitiques* , Payot, Paris, 1987

Calvet, Louis-Jean. *Pour une écologie des langues du monde*. Paris: Plon, 1999.

Guisan, Pierre. “O papel do discurso sobre a língua, como factor ou variável independente.” Assel – Universidade Federal Fluminense 2006

Mendes, Mafalda. “Cabo Verde - ir à escola em 12” in Revista da língua portuguesa e cooperação para o desenvolvimento (org) Mateus, Maria & Pereira, Luísa Teotónio. Língua Portuguesa e cooperação para o desenvolvimento. 2005 ed Edites colibri e CIDAC.

Monteiro-Lima, Adelaide – *A língua cabo-verdiana e a política linguística no país* – Cabo Verde . diss. Mestrado 2007. UFRJ - RJ

Construindo uma terminologia de raiz: textos legislativos sob exploração terminológica

Maria José Bocorny Finatto, Francine Facchin Esteves e
Guillermo Silva Villar

Sobre o autor

Maria José Bocorny Finatto (UFRGS, docente e pesquisadora do CNPq, mariafinatto@gmail.com)

Francine Facchin Esteves (UFRGS, mestranda em Letras, esteves.francine@gmail.com)

Guillermo Silva Villar (UFRGS, graduando em Letras, bolsista de IC, gsilvavillar@gmail.com)

Resumo

Este texto apresenta uma síntese de princípios e procedimentos básicos para quem deseja dar seus primeiros passos no trabalho de reconhecimento terminológico (RT) baseado em fontes escritas em formato digital. Essas fontes escritas servem como um *corpus*. *Corpus*, plural *corpora*, é o nome atribuído a um dado conjunto de textos, reunidos em formato digital, que se tenha interesse de examinar quanto ao vocabulário e/ou conteúdo veiculado. Tomam-se, como exemplos de *corpus/corpora*, alguns textos legislativos do Brasil, destacando-se: a) o texto do Estatuto da Criança e do Adolescente (ECA), conforme apresentado na Lei 8.069-90, promulgada em 13 de julho de 1990 e atualizada em 2021; b) o texto da nossa Constituição de 1988 (CF88). Em primeiro lugar, é feita uma breve introdução sobre as etapas mais básicas do RT, a partir de um ponto de vista linguístico e terminológico. Tratam-se da tipologia de textos e de discursos dos documentos-fonte, da seleção de fontes documentais e das distinções mais elementares entre palavras comuns e terminologias conforme estão empregadas em diferentes textos e cenários comunicativos. Em seguida, apresentam-se as funcionalidades da ferramenta AntConc, um recurso *on-line* de acesso gratuito e de fácil uso, que não exige conhecimentos avançados sobre Informática. Suas funcionalidades são ilustradas à medida que se explora o texto da CF88, cuja ponto para coleta inicial também é indicado. Dessa ferramenta, são ilustradas as funcionalidades: *wordlist* (lista de palavras); concordance (gerador de contextos por palavra ou expressão de busca); e *n-grams* ou *clusters* (grupos de palavras que se repetem ao longo de um texto ou arquivo). Apresentadas essas funcionalidades, faz-se um quadro dos achados sobre o vocabulário empregado na CF88, indicando-se alguns dados estatísticos sobre usos de palavras e terminologias mais e menos frequentes. Com isso, espera-se que o leitor consiga

compreender o modo básico de funcionamento da ferramenta AntConc. A partir desse auxílio, poderá identificar e organizar elementos úteis para panorama inicial do vocabulário e das terminologias empregadas ao longo da CF88 e de outros textos legislativos. Alerta-se, contudo, que cada sistema e ordenamento jurídico terão as suas próprias especificidades e tradições de expressão, peculiaridades que precisam ser reconhecidas na primeira fase do trabalho, etapa da compreensão e familiarização do terminólogo com a área de conhecimento com que se trabalha. Por fim, apontam-se algumas limitações e expansões possíveis a partir desse RT inicial.

Palavras-chaves

Reconhecimento Terminológico; Terminologia; ferramentas computacionais.

1. Introdução

Este texto traz uma síntese de conceitos fundamentais e dos procedimentos mais básicos para quem deseja se familiarizar com a realização de um reconhecimento terminológico (doravante RT), partindo de um conjunto documental escrito em formato de arquivo digital. Como exemplo de conjunto documental, utilizaremos alguns textos legislativos do Brasil, destacando a Constituição de 1988 (doravante CF88). Nesse sentido, vale alertar que apresentaremos informações imaginando, como leitor(a), uma pessoa totalmente inexperiente nesse tipo de tarefa, alguém interessado em dar seus primeiros passos no reconhecimento de terminologias técnicas ou científicas em português feito com apoio informatizado. No fim deste texto, indicamos algumas fontes para seu aprofundamento nesse tema.

Na sequência, este artigo tem a seguinte ordenação: em um primeiro bloco, mais teórico, temos uma brevíssima introdução sobre as etapas mais básicas do trabalho terminológico, que têm a ver com a condição “de raiz” mencionada no título. Tal condição pode ser assim resumida: “começando-se do zero”. Desse modo, no primeiro bloco, começamos com a identificação do perfil da área de conhecimento (ou domínio) envolvida, com a tipologia de textos e de discursos dos documentos-fonte e com o uso e/ou propósito que terá o RT, incluindo uma figura de seu provável destinatário. Depois, trazemos algumas distinções elementares entre palavras “comuns” e as terminologias técnico-científicas em meio a diferentes textos legislativos. No segundo bloco do texto, de caráter mais prático e ilustrativo, apresentamos a ferramenta AntConc (ANTHONY, 2012) e as suas funcionalidades. Com apoio de uma série de imagens e figuras, tentamos mostrar que essa ferramenta é um bom exemplo de recurso computacional de acesso público e gratuito.

Já antecipando um pouco mais esse segundo bloco, para tentar persuadir o nosso leitor a seguir nos acompanhando neste texto, frisamos que usar tal ferramenta demandará apenas que se possua um computador pessoal de configuração simples e/ou básica. Não é preciso ter conhecimentos específicos de Informática, tampouco se estar conectado à Internet durante seu uso ao longo de uma dada tarefa. Após sua instalação, o AntConc poderá ser utilizado para auxiliar a exploração de quaisquer conjuntos documentais de seu interesse.

2. Bloco 1

2.1. Entre termos e palavras e uma terminologia de raiz

Dito de um modo simples, um reconhecimento terminológico (RT) equivale à identificação e à sistematização de denominações respectivas a conceitos conforme utilizadas em um dado campo ou área do conhecimento, seja esse campo técnico, tecnológico ou científico. Geralmente, o RT envolve a produção de uma “lista” de nomes (termos) associados aos seus significados (conceitos). Além disso, junto de cada item da “lista”, teremos um conjunto de informações que nos permitam contextualizar e entender o seu uso ao longo de um conjunto de documentos escritos. Esses documentos, compondo um *corpus*, servem-nos como uma referência para a coleta dos itens dessa nossa “lista”. Assim, os documentos também nos ajudam a situar termos e conceitos, além de caracterizar o *modus operandi* e as convenções linguísticas da área de conhecimento em que se inserem.

Desse modo, em um RT, busca-se identificar e vincular um dado vocabulário – entendido, *grosso modo*, como um conjunto “especial” de palavras – que algumas vezes poder ser bastante peculiar –, às noções específicas que tal vocabulário corresponde. Um exemplo desse processo de RT ocorre quando se identifica a palavra EROSÃO conforme seja usada em textos de Geografia ou de Odontologia. No primeiro campo, temos, em geral, associado ao termo, a ideia de um processo de desgaste de solo. Mas, no cenário odontológico, o valor da mesma palavra corresponderá, em linhas gerais, ao processo de desgaste da estrutura de um dente, também chamado pelo termo EROSÃO DENTAL. Aqui temos um exemplo de um “termo técnico”, reparando-se que ele corresponde a duas noções e processos diferentes, em Geografia ou Odontologia, ainda que haja em comum uma ideia básica de “desgaste lento”.

Naturalmente, podemos considerar que a mesma palavra, EROSÃO, também poderia ser utilizada em cenários não especializados, científicos ou técnicos. Um exemplo desse uso mais comum, na nossa linguagem cotidiana, mas ainda associado à ideia de um “desgaste lento”, teremos em textos de opinião ou de jornais diários. Vejamos, a seguir, um trecho do texto intitulado “O que é o amor”. Foi apresentado em 14/02/2021, em uma coluna jornalística de opinião, publicada em Portugal. O texto foi produzido associando-se aos temas e às comemorações da data conhecida como “Dia dos Namorados”:

Apesar de toda esta torrente de “**erosão sentimental**”, que o amor acarreta e encerra na vida dos amantes, ele também é indubitavelmente comunhão, proteção, partilha, beleza, doçura, afeição, arrebatamento, satisfação e felicidade. (Fonte: observador.pt/opiniaio/o-que-e-o-amor/, acesso em 02/03/2022)

O nosso leitor deve se perguntar: afinal, o que difere o termo “técnico” da palavra comum? O que teria a ver **erosão sentimental** com **erosão dental** ou, mesmo, com a noção de erosão em Geografia ou Geociências? Essa pergunta complexa e, ao mesmo tempo encantadora, é enfrentada todos os dias por quem se dedica à Terminologia. Em tempo: vale já explicar que Terminologia (grafada com T maiúsculo inicial), para nós, é uma subárea da Linguística Aplicada, que é uma ciência que se coloca em aplicação.

A Terminologia e os terminólogos dedicam-se a estudar – descrever e compreender

– os diferentes fenômenos linguísticos da comunicação técnico-científica, em seus variados cenários. Mas, voltemos à pergunta, sobre diferenças de usos de **erosão**. Uma resposta inicial é: o que diferencia uma terminologia de uma palavra “comum” é, em primeiro plano, o seu ambiente comunicativo. E, repetindo a ideia de uma das maiores autoridades da nossa área da Terminologia (CABRÉ, 2005), podemos dizer: uma palavra não é um termo técnico-científico, ela **está** nessa condição em determinados contextos, que conferem a ela um significado “especial”. Esse significado ou modo de compreensão especial, chamaremos, *grosso modo*, de **conceito**.

Em sendo “especial”, esse significado “conceitual” corresponderá a uma dada unidade de conhecimento que foi socio-historicamente construída. Essa unidade se situará em um dado campo do saber e em um dado espaço cultural, sendo possível identificar, inclusive, suas transformações e abrangências – e compartilhamentos – ao longo do tempo, em uma cronologia. Quer dizer, o entendimento científico do conceito abrangido pelo termo **erosão**, conforme expresso em documentos no século XVIII, não será exatamente o mesmo daquele que temos hoje, no século XXI. E isso vale para todas as terminologias e conceituações técnico-científicas, que naturalmente progridem e se transformam.

O conceito que tomamos como um exemplo, em sua abrangência, tem a ver também com uma formação, saberes ou práticas e, fundamentalmente, com pessoas que têm objetivos associados a um saber-fazer ou saber-conhecer em dado campo. São essas pessoas que estabelecem e desenvolvem o conceito – e também a denominação a ele associada. Assim, seguindo com o nosso exemplo, vejamos uma conceituação de EROSÃO em Odontologia, conforme apresentada – e discutida – em uma revista científica atual dessa área:

A erosão dental descreve o resultado físico de uma perda patológica, crônica, localizada e assintomática dos tecidos dentais duros pelo ataque químico da superfície do dente por ácido e/ou quelante, sem o envolvimento de bactérias. (Fonte: http://revodontobvsalud.org/scielo.php?script=sci_arttext&pid=S1413-40122010000100015, acesso em 02/03/2022)

Conforme o trecho e o texto completo mostram, o conceito é expresso e utilizado por pessoas com uma dada formação, em um dado cenário comunicativo e em dado espaço-tempo do século XXI. Como já mencionado e vale sempre frisar, um conceito é objeto de ponderações e tenderá sempre à revisão e/ou questionamento, o que é necessário em todas as práticas científicas.

Assim, o uso de um termo e do respectivo conceito também corresponderá a um dado estágio de conhecimento da especialidade em que se insere, neste exemplo, a Odontologia. Quer dizer, conceitos e termos não são fixos, eles variam. A variação da forma e do conteúdo de um “termo-palavra”, inerente e natural à linguagem humana, sendo verificada pelo analista, precisará ser identificada e compreendida em um RT. E essa variação, seja pelas denominações, seja pelas conceituações, tende a se apresentar quando nos baseamos em um conjunto documental-fonte. Nesse conjunto, veremos as terminologias “ao vivo e a cores”, conforme usadas.

No cenário de saberes e denominações, desenhado e testemunhado por um conjunto de fontes documentais escritas, podemos dizer que as terminologias são personagens centrais ou muito importantes. Elas têm suas diferentes facetas, umas marcadas por denominações que se poderia reconhecer à primeira vista como “peculiares ou pouco usuais”, tal como, por exemplo, PERÓXIDO DE HIDROGÊNIO. Isto é, ao observar essa expressão, em

um texto, não duvidamos de que se trate de uma terminologia. Mas, por outro lado, há também aquelas terminologias que parecem estar disfarçadas de “palavras corriqueiras”, cujo significado acabou por tornar-se diferenciado quando foram alçadas a uma condição terminológica. Assim, temos palavras “comuns” que, nesse contexto diferenciado de saberes e conhecimentos, “viraram termos”.

Um exemplo disso seria a palavra CRIANÇA, muito corriqueira no nosso dia a dia. Como seu significado básico, via de regra, entendemos algo como “pessoa não-adulta”. Todavia, quando empregada e “significada” em um dado ambiente comunicativo de especialidade, essa palavra “comum” também assume contornos diferenciados. Um exemplo de tal transformação é o contexto jurídico do nosso *Estatuto da Criança e do Adolescente*, importante documento brasileiro conhecido como ECA, que corresponde à Lei 8.069-90, atualizada em 2021. Vejamos como se define/explica o que se entende por CRIANÇA nesse documento, uma fonte – ou *corpus* – que podemos enquadrar no domínio do Direito Civil do Brasil: “Art. 2º Considera-se **criança**, para os efeitos desta Lei, a pessoa até doze anos de idade incompletos, e adolescente aquela entre doze e dezoito anos de idade” (BRASIL, 1990).

Como se percebe, há um significado “especializado”, jurídico, uma delimitação em termos de anos de idade, que se soma ao nosso entendimento mais comum de **criança**. Aqui temos uma conceituação jurídica específica/particular associada a um dado termo que, à primeira vista, não parece ser um termo.

2.2. Reconhecimento terminológico: perseguindo palavras em transformação

Conforme vimos na seção anterior, tanto as palavras “comuns”, que empregamos no nosso dia a dia, e mesmo as palavras com um formato diferenciado, mais “técnico”, podem assumir condições de termo em um dado ambiente de significação. No caso do segmento de lei acima, o ECA, podemos considerar que há uma definição específica para CRIANÇA, que se opõe à de ADOLESCENTE.

Além disso, essa definição é circunscrita, isto é, ela vale apenas em um dado contexto, correspondendo a um valor que estabelece frente a todo um CONJUNTO DE OUTROS TERMOS E CONCEITOS com ela relacionados. Isso é o que chamamos de **sistema conceitual, que tem a ver como uma rede de conceitos e terminologias que percebemos pelo exame dos documentos de um RT**. Quando fazemos um RT, tentamos também capturar essas correlações entre as terminologias cujo uso identificamos em documentos-fonte. É como se as terminologias desenhasssem um “sistema”, que vamos, aos poucos, percebendo, conforme exploramos seus textos de ocorrência.

Entretanto, um texto não é feito somente de terminologias. No mesmo segmento do ECA, encontramos também palavras “comuns”, mas que dessa vez, seguem sendo assim, não associadas a um conceito ou a alguma especificação de sentido. Palavras desse tipo, que podemos considerar como típicos exemplos de não-termos, são, no trecho antes citado, o verbo CONSIDERAR, além das palavras PARA, OS, PESSOA, AQUELA e ANOS.

Ao realizarmos um reconhecimento terminológico (RT), vamos explorar um conjunto de textos que nos servem de referência e de espelhamento para uma dada área de conhecimento. Assim, será muito importante que esses textos-fonte sejam selecionados de acordo com uma boa condição de representatividade para a área em estudo. A partir daqui, vamos supor que

um RT envolve identificar, em diferentes documentos relevantes previamente selecionados, os seguintes elementos:

- a) TERMOS e seus respectivos CONCEITOS
- b) TERMOS e seus respectivos FORMATOS LINGÜÍSTICOS
- c) TERMOS, CONCEITOS e respectivos TERMOS E CONCEITOS RELACIONADOS.

Nos itens a) e b), acima, entra em jogo uma questão muito importante: **a variação terminológica**. Essa variação tem a ver com as diferentes formas das denominações, dentro de uma dada especialidade, que um TERMO pode ter. No caso do nosso primeiro exemplo, em Odontologia, já podemos assinalar um caso de variação denominativa, que pensamos ser de sinonímia: EROSÃO e EROSÃO DENTAL. Se nosso conjunto documental de referência incluir textos antigos, como seriam, por exemplo, estudos odontológicos do século XIX, a concepção associada ao termo pode ser outra. Isso é o que chamamos de **variação conceitual diacrônica** e se estuda em Terminologia Histórica¹.

Mas vamos pensar, por enquanto, em apenas textos da Odontologia da atualidade. Há, nos textos, o termo **erosão**, na sua forma resumida ou mais sintética, sem o adjetivo **dental**. Lendo o documento, poderíamos concluir que há um mesmo conceito associado às duas denominações – **erosão** e **erosão dental**. Isso nos mostra um caso de **variação denominativa sincrônica**, que ocorre em um mesmo espaço de tempo – e aqui em um mesmo documento-fonte.

Como se vê, os usos da linguagem, mesmo nesse contexto específico hipotético, também comportam fenômenos como sinonímias, sentidos diferentes e variações de escrita. Essas variabilidades, embora possam atrapalhar um pouco uma ideia e o ideal de precisão conceitual absoluta nas linguagens especializadas, apenas demonstram uma Humanidade envolvida nos saberes e conhecimentos. Nas ciências e nas técnicas, há pessoas envolvidas, pessoas que produzem os textos, cuja escrita espelhará a construção de uma dada matéria ou especialidade. Assim, ao começar um RT, você já lidará com textos, pessoas e seus saberes, o que, como analistas, nos cabe identificar e sistematizar. Fazemos isso em uma abordagem linguística dedicada, em especial, ao reconhecimento do vocabulário empregado.

Essa ideia de “vocabulário” corresponde, *grosso modo*, a uma realização do que os linguistas chamam de “léxico”. Mas, neste texto, não vamos entrar nessa distinção, que fica no cenário científico da Linguística.

Para administrar a variabilidade de termos e conceitos, sem a ideia de condená-la, pois o enfoque linguístico em um RT é sempre descritivo, é que temos os vocabulários controlados e/ou padronizados. Esses padrões de denominações geralmente são colocados pela autoridade de órgãos profissionais associados a uma dada especialidade. Nesses vocabulários, encontramos as “terminologias padronizadas” e também as “normas técnicas” de uma área. Assim, uma forma de denominar um respectivo conceito-significado é estabelecida em um dado contexto, de modo a se garantir precisão e boa correlação com outros termos e conceitos. Isso será importante especialmente em situações de trocas de conhecimento e de trocas em geral.

Vejamos o que seria um exemplo atual da compra, em larga escala, de toneladas de produtos lácteos: é preciso ficar bem claro, para comprador e fabricante-fornecedor, que

¹ Para saber mais sobre Terminologia Histórica, sugerimos: <<http://www.ufrgs.br/textecc/terminologia/>>.

se deseja, por exemplo, uma tonelada de **leite integral**, outra de **leite semidesnatado** ou, mesmo 5 toneladas de **leite UHT**. Cada um desses termos corresponde a um conceito e a preços – bem específicos, que não se pode confundir ou trocar ou alternar com os demais. Assim, desenha-se um cenário comunicativo em que termos e noções padronizados, muitas vezes associados também a códigos numéricos fixos, assumem um papel vital. Nesse cenário, fenômenos de variação precisam ser muito bem administrados e contornados.

Guardadas as devidas diferenças, é semelhante o caso, por exemplo, do conceito de CRIANÇA frente ao conceito de ADOLESCENTE no nosso ECA. Crianças não poderão ser confundidas, em um cenário legal e jurídico, com adolescentes ou pessoas adultas, salvo condições especiais definidas naquele texto, que funciona como uma moldura de significação para suas terminologias. O mesmo vemos nos casos dos nomes “oficiais” para algumas doenças, que inclusive correspondem a um código numérico, conhecido como CID ou Classificação Internacional de Doenças. A ideia, nesse contexto de padronização das terminologias da área da Saúde, é evitar confusões e tentar garantir que todos possam ter um mesmo entendimento – ou conceito uniforme – de um dado TERMO + CONCEITO/ DESCRIÇÃO DE SEU SIGNIFICADO. Abaixo, alguns exemplos dessa padronização da CID para o termo SARAMPO e seus tipos.

CID 10 – B05 – Sarampo

CID 10 – B05.0 – Sarampo complicado por encefalite

CID 10 – B05.1 – Sarampo complicado por meningite

Feitas essas explicações sobre peculiaridades das terminologias, em suas diferentes circunstâncias e variabilidades de uso e de significações, passemos ao tema do reconhecimento terminológico (doravante RT) de um modo mais pontual. A partir daqui, podemos compreendê-lo como um tipo de trabalho de **mediação de comunicação** realizado por profissionais de uma área, terminólogos, linguistas, entre outros. Salienta-se, assim, a ideia de uma mediação terminológica (CONCEIÇÃO; ZANOLA, 2020).

Esse tipo de trabalho profissional, que se concretiza em um RT, como já citado, visa repertoriar nomes e conceitos e suas correlações. Esse repertório, como no caso das doenças listadas na CID, pode servir para ajudar diferentes pessoas a entenderem os significados “peculiares” ou “específicos” de um dado contexto de conhecimentos. Nesse exemplo, ainda que se trate de um texto-lista, o mediador de comunicação se depara com diferentes terminologias associadas à doença conhecida, genericamente, como **sarampo**.

Em geral, usamos fontes textuais escritas para o reconhecimento conceitual e denominativo em uma dada área. Isso porque, afinal, “o texto é o habitat das terminologias” (KRIEGER; FINATTO, 2004) e é ele que nos mostrará, por exemplo, que termos podem ocorrer, ligados a um mesmo conceito, com formas diferentes, como vimos em EROSÃO e EROSÃO DENTAL.

Ter um corpus documental é hoje uma condição *sine qua non* para um trabalho terminológico de qualidade, pois ele ficará amparado em fontes primárias que são um índice de confiabilidade da informação que se traz em um RT. Naturalmente, poderemos associar ao nosso *corpus* algumas ou várias fontes secundárias, como são as terminologias padronizadas disponíveis em catálogos ou “listagens” produzidos por grupos profissionais ou instituições; os diferentes dicionários preexistentes publicados por fontes avalizadas; os dicionários “comuns de língua”, feitos por lexicógrafos e estudiosos da Linguística; e as

bases de dados diversas que hoje o mundo da Internet nos oferece. Os dicionários comuns, embora sejam dicionários não-técnicos, são também muito úteis, pois podem nos ajudar, como terminólogos, a entender um significado básico – ou preliminar – de uma palavra-termo com a qual nos deparamos

O tipo de estudo linguístico que mais tem lidado com reconhecimento de vocabulários – de diferentes tipos – a partir de bases textuais em formato digital é a Linguística de Corpus (BERBER SARDINHA, 2000). Nesse tipo de Linguística, tem sido comum o desenvolvimento de métodos e ferramentas informatizadas para apoiar os pesquisadores, já que um dos seus princípios é o estudo o mais extensivo possível de diferentes acervos em formato digital, chamados de *corpus/córpus/corpora* (plural).

A Linguística de Corpus (LC) pode ser entendida como uma concepção teórica sobre o funcionamento da linguagem humana e também como uma metodologia para o seu estudo. Como concepção, entre várias, destacamos o entendimento que “uma palavra se define por suas companhias”. Isto é, em LC, entende-se, por princípio científico, que o sentido e o valor de uma palavra somente serão apreendidos a partir do entendimento, o mais amplo possível, de seus diferentes usos e contextos ao longo de diferentes *corpora* documentais que representam um dado estado de língua.

Entre os materiais que podemos usar para fazer um RT baseado em *corpus* e inspirado pela LC, podemos pensar que há diferentes práticas textuais associadas a uma dada especialidade ou área de conhecimento. Uma divisão entre fontes mais e menos especializadas, esclarecemos, não implica qualquer apreciação quanto à validade ou importância das terminologias empregadas nos documentos ou dos próprios textos-fonte. É somente uma referência inicial, um recurso teórico-metodológico para auxiliar nossa compreensão sobre a distribuição de conceitos e das terminologias ao longo de um dado acervo textual, que pode ter uma composição heterogênea. Assim, definir um tipo de documento ou de fonte para se lidar em um RT é fundamental.

Em Odontologia, temos como fontes, por exemplo, os artigos científicos, escritos por especialistas para especialistas, os livros ou manuais de estudo para a formação universitária de cirurgiões dentistas e também os textos de divulgação produzidos por esses especialistas para pessoas leigas interessadas, por exemplo, em como ter uma boa saúde bucal. No caso de um RT voltado para pessoas leigas, pessoas que vivam no Brasil, sem formação em Odontologia, interessadas apenas em conhecer as terminologias e conceitos mais recorrentes sobre o **câncer de boca** [ou **cancro de boca** – fora do Brasil], podemos pensar em **não usar**, como *corpus*, manuais de estudo universitário. Entretanto, a escolha documental vai depender do tipo de RT que se realiza e o objetivo que se tenha com ele.

Na área legislativa, no Brasil, também como exemplo dessa diversidade de fontes, há os textos de Doutrina e os textos da Jurisprudência, que são também de vários tipos: Decretos, Leis, Portarias etc. Ao empreendermos um RT no âmbito legislativo, logo após fixar a área ou domínio com que trabalharemos, devemos identificar que tipos de documentos melhor podem servir para os nossos propósitos. Assim, por exemplo, poderíamos imaginar um reconhecimento da terminologia do Direito Ambiental brasileiro, para o qual tomaríamos como referência todo um conjunto de leis, decretos e portarias válidos no nosso país em um dado recorte temporal. Esse reconhecimento, uma vez feito, pode abastecer um dicionário, tal como um que já temos (KRIEGER et al, 1998), ou mesmo alimentar uma base de dados para acesso on-line (www.ufrgs.br/termisul).

Um tal RT legislativo também poderia servir de apoio para um recurso didático voltado para o cidadão comum, sem formação em Direito, ou mesmo para diferentes estudantes universitários interessados na legislação ambiental do Brasil. Nesse caso, vamos imaginar um conjunto composto, por exemplo, por 800 leis, as quais versam sobre diferentes aspectos ambientais. Vamos supor que estamos trabalhando em um RT para uso de jornalistas que lidam com temas ambientais. Como explorar essas 800 leis para chegar, por exemplo, a um conjunto de seus termos e conceitos conforme sejam mais comumente empregados nessas leis? Como apresentar a informação de forma a melhor atender o nosso suposto usuário jornalista, que, sem ter formação em Direito ou Biologia, precisaria ler e entender a legislação?

Naturalmente, hoje, dada a larga prática de digitalização desse tipo de documento e a garantia de seu acesso a todo o cidadão, parece ser fácil encontrar e percorrer uma base de dados com leis ambientais. O Senado Federal do Brasil, por exemplo, oferece todo um banco de leis, decretos e outros documentos afins para acesso público. Basta a pessoa acessar um site determinado e salvar os documentos no seu computador. Feito isso, “bastaria” a pessoa – o jornalista que imaginamos – ler, calmamente e com cuidado, todas as 800 leis do nosso caso imaginário e ir fazendo um registro, em um arquivo de texto, de suas terminologias e conceituações à medida que avance com a leitura. Outra opção seria o “nosso” jornalista consultar um dicionário especializado sobre esse tema, mas é nele, como um ponto final possível, entre outros, de um RT que queremos chegar com o que tratamos neste texto.

Pois, no próximo bloco deste texto, vamos apresentar um método e um recurso – uma ferramenta informatizada – bastante úteis para o nosso explorador das terminologias empregadas nessas imaginárias 800 leis sobre Direito Ambiental e que estaria realizando um RT para atender as necessidades de jornalistas. A ferramenta, adiante apresentada, permite que se administrem, ao mesmo tempo, centenas de documentos captados em formato digital e que se gerem muitos registros, que podem ser armazenados de um modo organizado e, principalmente, recuperável.

Essa ferramenta, como quaisquer outras do tipo, frisamos, nunca substituirá o trabalho humano de leitura dos textos, da nossa ponderação sobre as tipologias de documentos e sobre as condições e necessidades envolvidas. Tampouco nos exime, como terminólogos, do trabalho de nos inteirarmos das práticas e modos de dizer específicos da área de conhecimento em foco, com seus valores, histórias e modos de compreender a realidade de seus especialistas.

Dito isso, vale voltar, ainda, às diferentes concepções de EROSÃO em Geografia, Geologia e em Odontologia ou mesmo às concepções de CRIANÇA ou mesmo de MULHER ou, ainda, de palavras como ABORTO, em diferentes cenários legislativos e jurídicos, extremamente marcados por condições sociais, históricas e políticas. Assim, reconhecer-se uma área de conhecimento, sua história, valores, estrutura e organização, incluindo suas práticas de escrita e comunicação, é um trabalho prévio e vital em qualquer tipo de RT. Lidando-se com leis e com a esfera jurídica isso parece ficar acentuado.

O terminólogo, portanto, antes de começar um RT lidando com dados que vai extrair e sistematizar de um *corpus* documental que toma como referência, deve aprender o máximo possível do domínio com o qual vai trabalhar. E isso é feito com um trabalho ainda mais anterior à seleção e recolha de textos do seu *corpus*. Esse é o trabalho que, frisamos, envolve saber para quem e para qual finalidade o RT está sendo feito. É para pessoas leigas na área? Qual o perfil dessas pessoas leigas? Qual o nível de escolaridade e que tipo de conhecimento

prévios tais leigos possuem? Ou será um trabalho para semiespecialistas, estudantes da área? O RT vai abastecer um dicionário tradicional ou uma base de dados para fins didáticos?

Todas essas perguntas são importante, porque nos ajudam a decidir com que tipos de documentos e de fontes textuais vamos trabalhar. Esse passo primordial do trabalho terminológico também permite estimar os recursos com que se precisará contar para chegar-se a um produto final de qualidade. Feito isso, passaríamos a uma caracterização de domínio e de fontes textuais, tendo em mente um RT para jornalistas que fosse dedicado, por exemplo, à legislação ambiental.

2.3. Algumas questões e peculiaridades linguísticas sobre os textos legislativos e legais do Brasil

O Direito brasileiro, é importante ressaltar, compõe-se de diferentes tipos de textos legais. Se em um primeiro passo seria cabível pensar que tudo estaria resumido no tópico “lei”. Mas, em um olhar mais detido, perceberemos que não é o caso. Podemos, assim, apontar a situação hierárquica elencada pela própria estrutura do nosso “ordenamento jurídico” (expressão que é também uma terminologia). Nesse sentido, caberia já pontuar a prevalência da Constituição como objeto primário do qual partiria toda a construção jurídica do país. No entanto, como mencionado, não é apenas nas leis que repousa a estruturação do sistema jurídico. Além delas, podemos citar os elementos doutrinários e a jurisprudência decorrente da aplicação jurídica.

Da mesma forma, é possível perceber que, em uma análise linguística, também se nota uma formalização da linguagem aplicada de forma sistêmica. Tomemos como exemplo o artigo 205 da CF88, que dispõe sobre inúmeros direitos:

Art. 205. A **educação**, direito de todos e dever do Estado e da família, será promovida e incentivada com a colaboração da sociedade, visando ao pleno desenvolvimento da pessoa, seu preparo para o exercício da cidadania e sua qualificação para o trabalho. (BRASIL, 1988; grifo nosso)

Vale perceber que a linguagem empregada nesse trecho do documento tenta, a seu modo, ser de “uso corrente”, mas nela estão naturalmente inseridos modos de dizer e elementos (conceitos) de cunho mais restritivo do que aqueles adotados livremente em um diálogo informal. Como exemplos, temos a menção de “direito” e “Estado”, termos “técnicos” que consubstanciam a intrincada labuta da compreensão jurídica. Agora observemos um desdobramento legal que pode ser encontrado no ECA (Estatuto da Criança e do Adolescente), que versa sobre o mesmo assunto, a educação:

Art. 53. A criança e o adolescente têm direito à **educação**, visando ao pleno desenvolvimento de sua pessoa, preparo para o exercício da cidadania e qualificação para o trabalho, assegurando-se-lhes:

- I – igualdade de condições para o acesso e permanência na escola;
- II – direito de ser respeitado por seus educadores;
- III – direito de contestar critérios avaliativos, podendo recorrer às instâncias escolares superiores;
- IV – direito de organização e participação em entidades estudantis;
- V – acesso à escola pública e gratuita próxima de sua residência.

Parágrafo único. É direito dos pais ou responsáveis ter ciência do processo pedagógico, bem como participar da definição das propostas educacionais (BRASIL, 1990).

Cabe aqui ponderar dois aspectos, primeiramente, note-se que há uma clarificação do que consiste o direito à educação citado pela CF88. Diante disso, compreende-se a razão deste tratar de uma norma infraconstitucional, ou seja, decorrente da Constituição, que fica “abaixo” dela. A mesma subordinação valeria, por exemplo, para uma lei ambiental que discipline o transporte de **cargas perigosas** ou que tratasse sobre o, sempre pungente, tema dos **agrotóxicos** ou **pesticidas**.

Em outro ponto, quanto à forma que a linguagem usualmente assume no texto jurídico e/ou legal, ainda que se possam vislumbrar tentativas de formulações mais compreensíveis para as pessoas “comuns”, há muito o que explorar no tocante à sua “leiturabilidade” ou acessibilidade textual e terminológica (ATT) (FINATTO, 2020).

A propósito da ATT, sugerimos que o nosso leitor tenha acesso aos apontamentos de Motta (2021), cujos estudos tratam sobre as terminologias jurídicas e sobre a sua compreensibilidade maior ou menor para pessoas leigas. Uma tal compreensão variável, vale dizer, maior ou menor, pode estender-se também a um terminólogo pouco familiarizado com a linguagem jurídica e suas práticas textuais em meio às suas tarefas para um RT. Essa autora tem investigado padrões do texto de sentenças judiciais, justamente, com o apoio de ferramentas computacionais e da LC. Tem verificado que, para além das terminologias, há todo um emprego de palavras e de construções frasais muito recorrentes. Esse emprego, se comparado com o que verifica em *corpora* jornalísticos, assinala pontos importantes quando se pensa em dificuldades de compreensão de leitura.

Perceber, quantificar e analisar fenômenos e características do uso das linguagens especializadas, para além de um RT, serve também para descrição sistemática dessas linguagens. Assim, um dado tipo de texto e de discurso, integrando um *corpus* ou *corpora*, ao ser explorado com apoio informatizado, oferecerá outras perspectivas sobre conhecimentos, saberes, culturas e práticas de escrita.

3. Bloco 2

3.1. Conhecendo uma ferramenta útil de acesso gratuito

A Linguística de Corpus (LC), conforme já mencionamos, envolve o estudo da linguagem por meio da análise de conjuntos de dados linguísticos compilados, em formato eletrônico, de acordo com critérios cuidadosamente predefinidos pelo pesquisador. Mediante os dados textuais processados com auxílio computacional, a LC, como metodologia, tem apontado alternativas para que pesquisadores observem, de um modo extensivo, padrões e recorrências da linguagem em um enfoque quantitativo e qualitativo.

Um dos produtos oriundos da análise de *corpus/corpora*, com esse auxílio, pode ser um RT ou mesmo um glossário ou dicionário de dado campo de especialidade. Em linhas gerais, podemos dizer que as ferramentas de LC nos permitem obter, de modo quase instantâneo, diferentes tipos de listas de usos e de frequências de palavras ou de terminologias conforme empregadas ao longo de toda uma dada coleção documental ou de várias coleções, ao mesmo tempo.

Entre as ferramentas utilizadas em pesquisas desse tipo, o AntConc (ANTHONY, 2012) é uma ferramenta multiplataforma para investigações linguísticas em LC. É muito utilizada para analisar e fazer pesquisas, simultâneas, em muitos ou em poucos textos, ou mesmo em apenas um texto, como seria um texto de lei, ou todo um conjunto de leis que nos interesse explorar para um RT. Essa ferramenta foi desenvolvida pelo Professor Dr. Lawrence Anthony, docente e pesquisador linguista, que atua em uma universidade do Japão. Essa ferramenta pode ser baixada, no computador pessoal do nosso leitor, gratuitamente, em <https://www.laurenceanthony.net/software/antconc/>.

Por ser de alta compatibilidade com os sistemas operacionais atuais, o AntConc acaba sendo uma opção viável para análises em corpus em português feitas em diferentes tipos de textos e de linguagens. Seu único senão, na nossa opinião, é que sua interface de uso está apenas em inglês. Mas, para quem tenha alguma dificuldade, mesmo após ler este nosso texto até o final, há diferentes tutoriais e guias de uso sobre o AntConc em português em diferentes fontes (exemplo: <https://www.youtube.com/watch?v=nyPrxMTvHxI>, que acessamos em março de 2022).

A seguir, temos um passo a passo para utilizar a ferramenta AntConc. Mas, em primeiro lugar, antes de tudo, é preciso ter um texto ou vários textos ou um *corpus* de seu interesse para percorrer com essa ferramenta. Essa fonte documental será aqui tomada como uma referência que o nosso leitor poderá expandir conforme for do seu interesse e/ou conforme tenha se familiarizado com as suas funcionalidades.

a) Passo 1: seleção de um *corpus* e/ou texto como fonte do RT

Começamos com o ponto primordial, a escolha do *corpus* a ser utilizado. Vamos voltar a imaginar aquele RT dedicado ao Direito e à Legislação Ambiental do Brasil, que incluirá várias leis, mas também a lei-base do país, que é a sua constituição, a nossa CF88. Iniciaríamos, então, pela CF88, texto facilmente encontrável na Internet em vários formatos: Imagem, PDF, HTML, somente texto (.txt) e doc (.doc ou .docx). Conforme o tipo de arquivo com que se trabalhe, será preciso realizar ou não alguns procedimentos antes do seu processamento pela ferramenta AntConc.

Por uma questão de praticidade de formatos de arquivo, optamos por mostrar ao nosso leitor uma sua versão já pronta para uso, conforme oferecida no site do grupo TERMISUL. O TERMISUL² é um dos grupos acadêmicos mais antigos do Brasil dedicado à pesquisa e à formação em Terminologia. Seu *site* oferece *corpora*, diferentes dados, materiais de estudo e diferentes insumos úteis para quem trabalha com RTs em diferentes idiomas, especialmente com o português.

Uma vez conhecido esse *site*, no seu todo, vá diretamente para a página <http://www.ufrgs.br/termisul/paises.php>. Nela, você percorrerá uma lista com nomes de países até encontrar “Brasil”. Veja que, neste ponto, você encontra os itens “Caracterização” e, ao lado, “Baixar”. No primeiro, há uma síntese sobre a estrutura e o tipo de linguagem encontrado nos documentos legais oferecidos. No segundo, o arquivo da CF88, conforme figura a seguir ilustra.

²TERMISUL (Projeto Terminológico Cone Sul). Dados gerais em: www.ufrgs.br/termisul. Especificamente, sobre o grupo de pesquisa, veja em: <http://dgp.cnpq.br/dgp/espelhogrupo/6996455967777098>

IMAGEM 1

Página do site TERMISUL onde se encontra o arquivo da CF88 e uma listagem alfabética de documentos por nomes de países.



Esse tipo de informações, a Caracterização, para além do arquivo de texto que você vai usar, pode ser muito valioso. Afinal, nos ajuda em um reconhecimento prévio do texto e do discurso associados ao documento que nos interessa explorar com a ferramenta AntConc. Veja um exemplo do tipo de informação oferecida pela equipe TERMISUL sobre a CF 88:

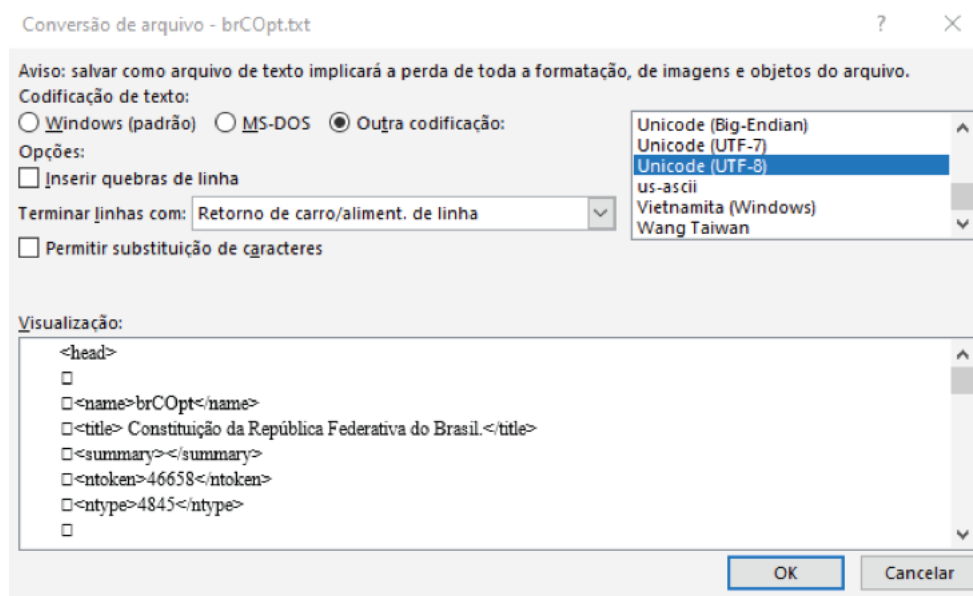
O texto apresenta preâmbulo, corpo do texto, local, data e assinaturas. O preâmbulo contém a fórmula da promulgação: "Nós, representantes do povo brasileiro, reunidos em Assembléia Nacional Constituinte para instituir um Estado Democrático, [...] promulgamos, sob a proteção de Deus, a seguinte CONSTITUIÇÃO DA REPÚBLICA FEDERATIVA DO BRASIL."

O corpo do texto se organiza em nove títulos, subdivididos em capítulos, muitos desses em seções, perfazendo um total de 250 artigos e ainda os atos das disposições constitucionais transitórias, formados por 94 artigos. Muitos artigos são subdivididos em incisos, parágrafos e alíneas. (TERMISUL)

O arquivo baixado do *link* acima provém do *corpus* legislativo TERMISUL. Ele estará em formato .DOC, geralmente utilizado pelo editor de texto Word ou outro. Esse ponto já antecipa uma ação preliminar fundamental para a composição do seu *corpus* e alimentação da ferramenta AntConc: **a conversão do arquivo .doc para o formato .txt.**

Dentro do seu próprio editor de texto, o texto da CF88 que você baixou é um arquivo comum, como os arquivos que você costuma usar. Mas, para poder usar o texto na ferramenta, será necessário ir em "Salvar como" e optar por "Texto sem formatação (.txt)". É importante especificar "Unicode (UTF-8)", pois **somente assim** o AntConc conseguirá "ler" acentos, cedilhas, til e outros caracteres ao longo do seu texto-arquivo da CF88. Veja, na figura abaixo, uma etapa desse procedimento "Salvar como".

IMAGEM 2
Conversão de arquivo .doc para .txt no Word



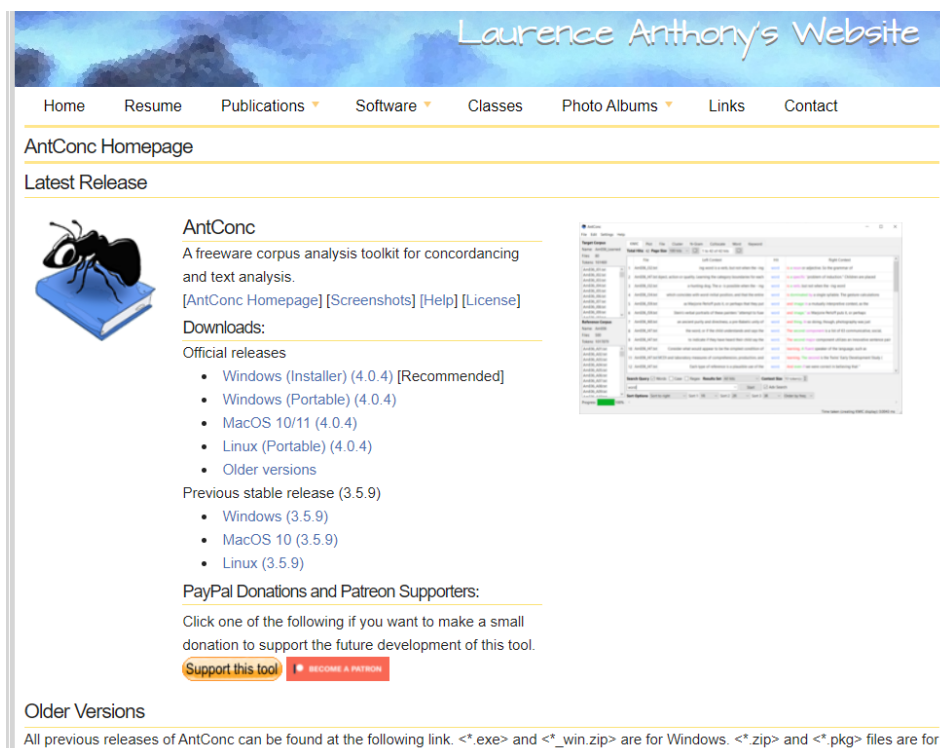
Se você, depois de conhecer o que mostramos aqui, for trabalhar com um arquivo que obteve apenas em formato PDF, deverá, obrigatoriamente, também transformá-lo em um arquivo de texto no formato .txt. Para esse caso, siga este breve passo a passo usando um conversor automático de PDF para .doc (https://www.ilovepdf.com/pt/pdf_para_word) ou simplesmente abrindo o arquivo PDF no seu editor de texto.

b) Passo 2: *download* do AntConc

Agora que você já tem seu arquivo em formato .txt, abra o arquivo e veja se está tudo ali, conforme se indica na Caracterização no site do TERMISUL. Estando tudo de acordo, você terá certeza de que tem o texto completo para analisar. Então, acesse o site da ferramenta (como mencionado, com interface e instruções apenas disponíveis em inglês): <https://www.laurenceanthony.net/software/antconc/>.

Como você consegue ver na Figura 3, será possível escolher o modelo de programa mais adequado para o seu sistema operacional, ou seja, Windows, Linux ou Mac. No site, também será possível encontrar outras informações sobre a ferramenta AntConc, como uma seção destinada para o auxílio (HELP) que aborda tanto o processo de *download* e escolha de qual modelo melhor se adequa, assim como o uso da ferramenta em si. Há também uma área destinada a responder as perguntas mais frequentes, o que pode “trazer luz” para alguns usuários, em caso de necessidade ou dúvida. Neste texto, tratamos da versão 3.5.9, pois a versão mais recente (4.0.4), lançada em 2022, ainda apresentava instabilidades.

IMAGEM 3
Página inicial do site da ferramenta AntConc.

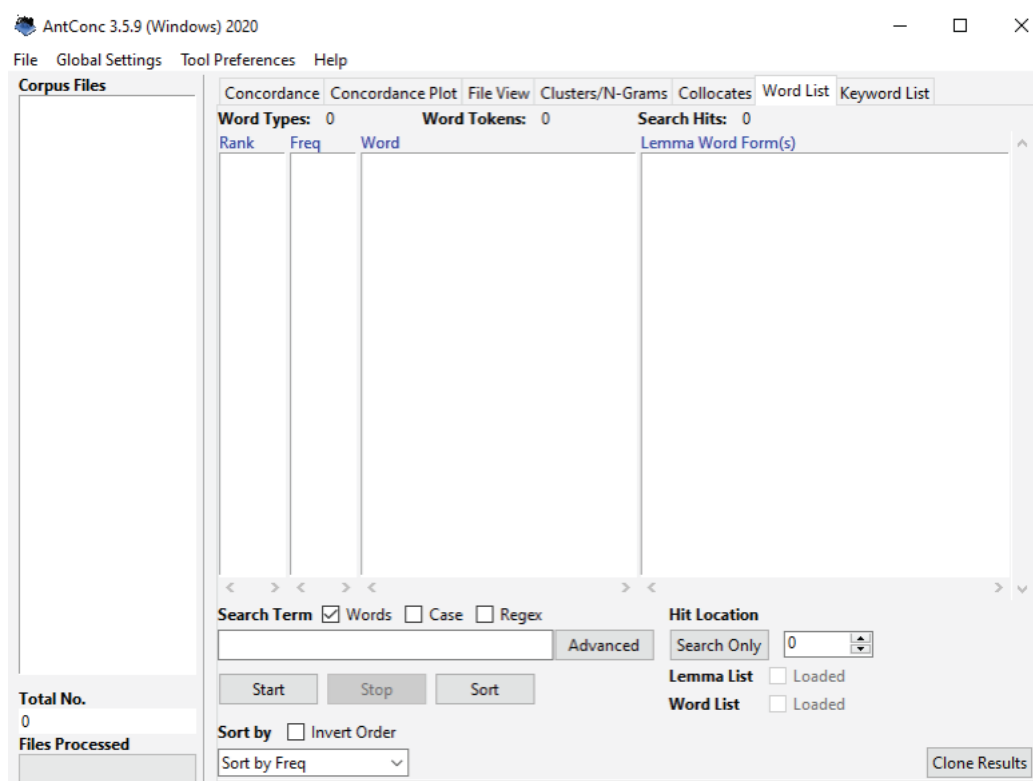


c) Passo 3: reconhecendo a interface do AntConc

Após a instalação da ferramenta no seu computador, é possível ver o ícone do AntConc em seu computador que tem o desenho de uma formiga em cima de um livro. Basta clicar duas vezes nesse ícone para começar a usar.

Para carregar um corpus/texto a ser explorado pela ferramenta, o usuário deve clicar em "File" (arquivo) e selecionar a primeira opção "Open file(s)" (abrir arquivos). Uma janela será aberta e nela, o usuário pode navegar até encontrar o *corpus* que pretende carregar na ferramenta, lembrando que os arquivos precisam estar em formato .txt e previamente salvos na opção UNICODE UTF-8. É possível perceber na Figura 4 que existem várias opções para serem utilizadas dentro da ferramenta. Abordaremos as principais mais adiante.

IMAGEM 4
Interface da ferramenta AntConc

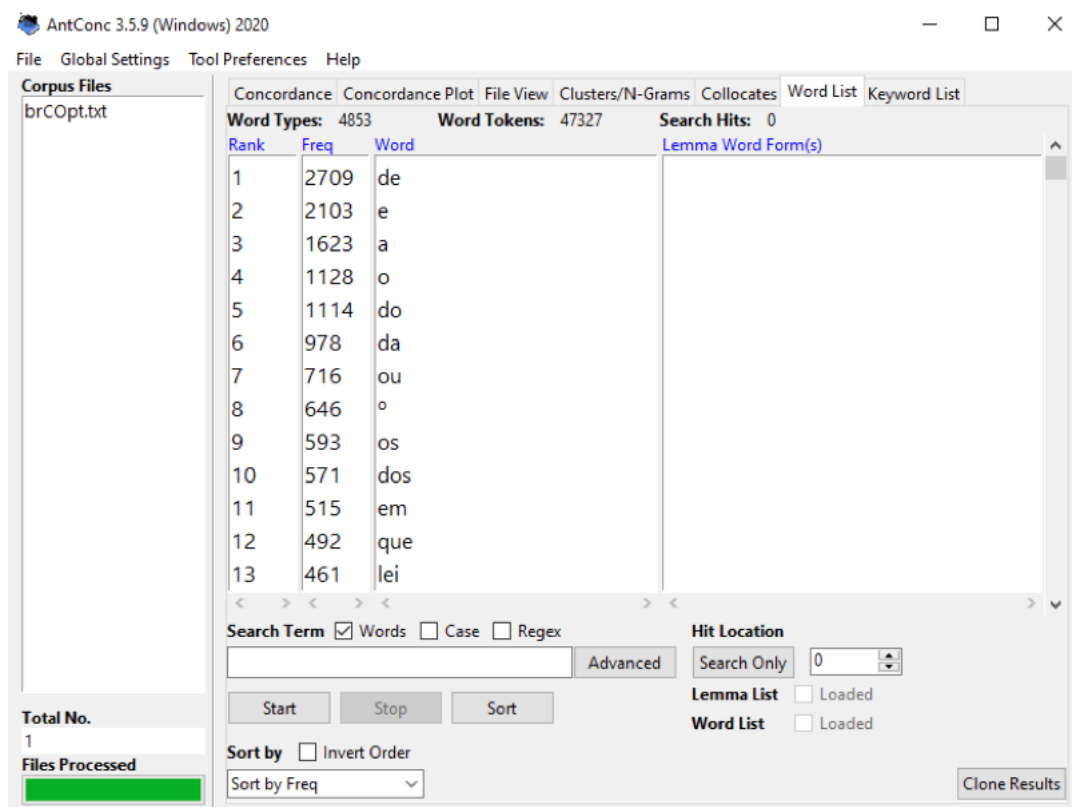


As opções de análise e geração de listas encontram-se na parte superior ao centro e variam em escopo, resultando, portanto, em formas diferentes de se visualizar o texto. O texto inserido na ferramenta se visualiza em *File View*. As funcionalidades mais utilizadas e recomendadas em um RT inicial seriam: *Wordlist*, *Clusters/N-Grams* e *Concordance*. É válido ressaltar que a nomenclatura de algumas dessas funções sofreu alteração ao longo do tempo, de tal forma que, dependendo da versão do sistema que você baixou no seu computador, algumas podem aparecer com nomes diferentes, ainda que tenham o mesmo objetivo.

I. *Wordlist* (lista de palavras)

Essa lista enumera e agrupa (se repetidas) todas as palavras do *corpus* analisado, podendo ser ordenada de forma alfabética, por frequência etc. Nela temos as seguintes informações: ordem, frequência, palavra. A ordem representa o lugar que tal palavra encontra-se dentro do critério de organização da lista; a frequência, quantas vezes a palavra ocorre no *corpus* analisado; e, por fim, a palavra em si (Figura 5).

IMAGEM 5
Aba Wordlist da ferramenta AntConc



Veja que aparecem no alto, em cima da lista, as informações numéricas sobre **types** e **tokens**. Os **types** correspondem ao número de *palavras diferentes* que há no seu texto/arquivo. Já os **tokens** mostram o número total de palavras que o AntConc contou no texto da nossa CF88.

Com esses dois números, usando uma calculadora simples, ao dividir o número de **types**, *multiplicado por 100*, pelo número de **tokens**, você obtém uma razão/proporção chamada de TTR, que é o percentual de palavras diferentes em relação ao número total de palavras presentes no texto. Quanto maior for a proporção de **types**, frente ao total de palavras de um texto, mais variado será, em tese, o seu vocabulário. Quando essa conta tem 100% como resultado, temos que o texto é todo feito somente de palavras que nunca se repetem, como uma lista de compras. Essa contagem, entre outras, já poderia dar uma ideia do perfil de vocabulário de um texto legislativo como a CF88. Seria ele mais variado ou mais repetitivo frente ao de outros tipos de documentos?

Um exemplo: um artigo científico, escrito em português, que tende a cultivar a repetição de termos e de palavras, para facilitar a replicabilidade dos experimentos que descreve, tende a mostrar uma TTR em torno de 30%. Para ponderar o valor da TTR na nossa CF (com resultado em torno de 10% [4.853 palavras diferentes vezes 100 dividido pelo número total de palavras – 47.327], podemos considerar os seguintes exemplos de valores médios de TTR para outros tipos de textos. Entre os quais, a CF88 seria um texto com o menor valor proporcional – o texto com menor o valor de TTR (10%), sendo que traz o vocabulário, proporcionalmente, mais repetitivo entre todos. Vejamos:

- um artigo científico de Biologia em português (29%);
- uma redação de vestibular da UFRGS com nota excelente (37,14%);
- uma notícia de um jornal diário, como o ZH, de Porto Alegre/RS (59,6%). Aqui cabe ressaltar que a escrita jornalística tende a evitar, ao máximo, as repetições de palavras.

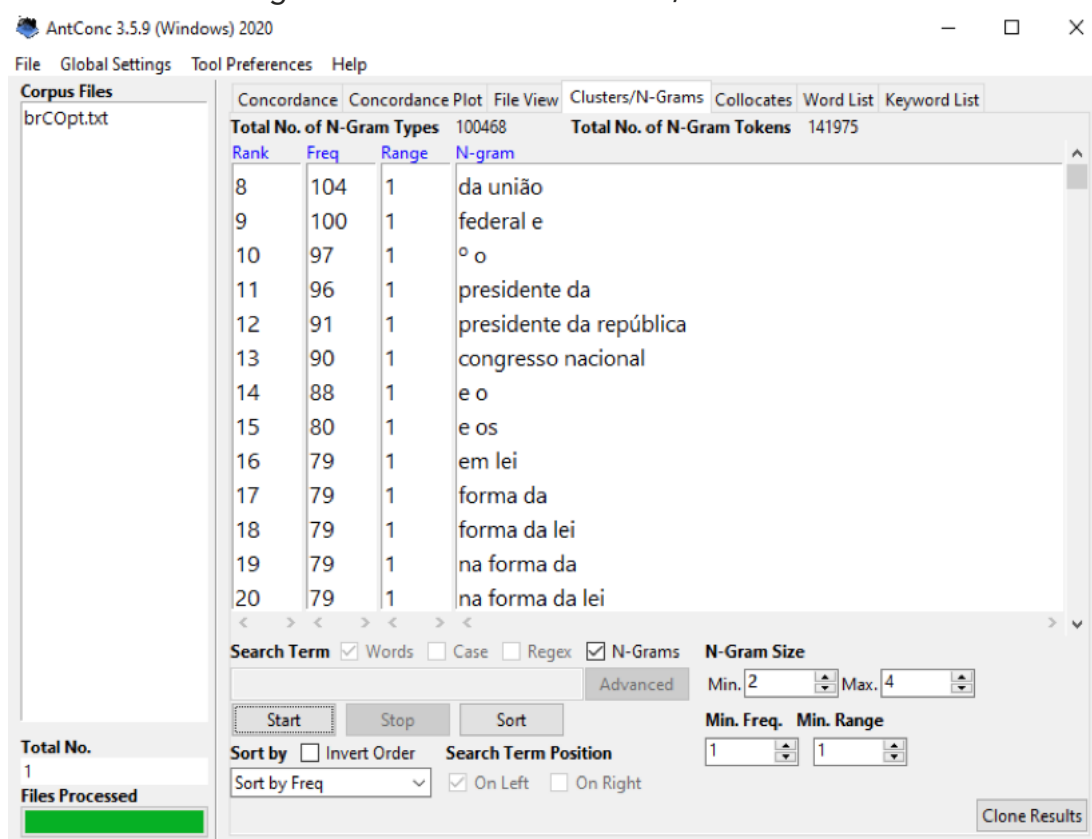
Assim, nesse quadro comparativo, à primeira vista, o texto da CF88 aparece como uma fonte escrita que teria um vocabulário “super” repetitivo. Essa característica, de antemão, já poderia sinalizar algumas consequências para o nosso RT.

II. N-grams (lista de palavras que aparecem agrupadas)

A aba *N-Grams* opera com um conjunto de palavras agrupadas, cuja quantidade pode ser arbitrada pelo usuário para verificar expressões recorrentes e junturas de palavras específicas no texto. No exemplo elaborado, em que selecionamos o tamanho do n-grama como no mínimo 2 e no máximo 4, podemos ver, na Figura 6, que a lista conta com três conjuntos de palavras: de duas palavras (bigramas), de três palavras (trigramas) e de quatro palavras (quadrigramas). Esses agrupamentos de palavras podem ser classificados como combinatórias ou fraseologias, podendo ter valor terminológico ou não. Para além da lista de palavras gerada na aba *Wordlist*, os agrupamentos de palavras também merecem sua análise e descrição e podem ser incluídos em obras terminológicas. Por exemplo, os termos “pessoa física” (7 ocorrências), “pessoa jurídica” (8 ocorrências) e “pessoa física ou jurídica” (6 ocorrências) têm valor terminológico justamente mediante a junção de palavras no contexto jurídico.

IMAGEM 6

Aba N-grams da ferramenta AntConc, com tamanho de 2 a 4

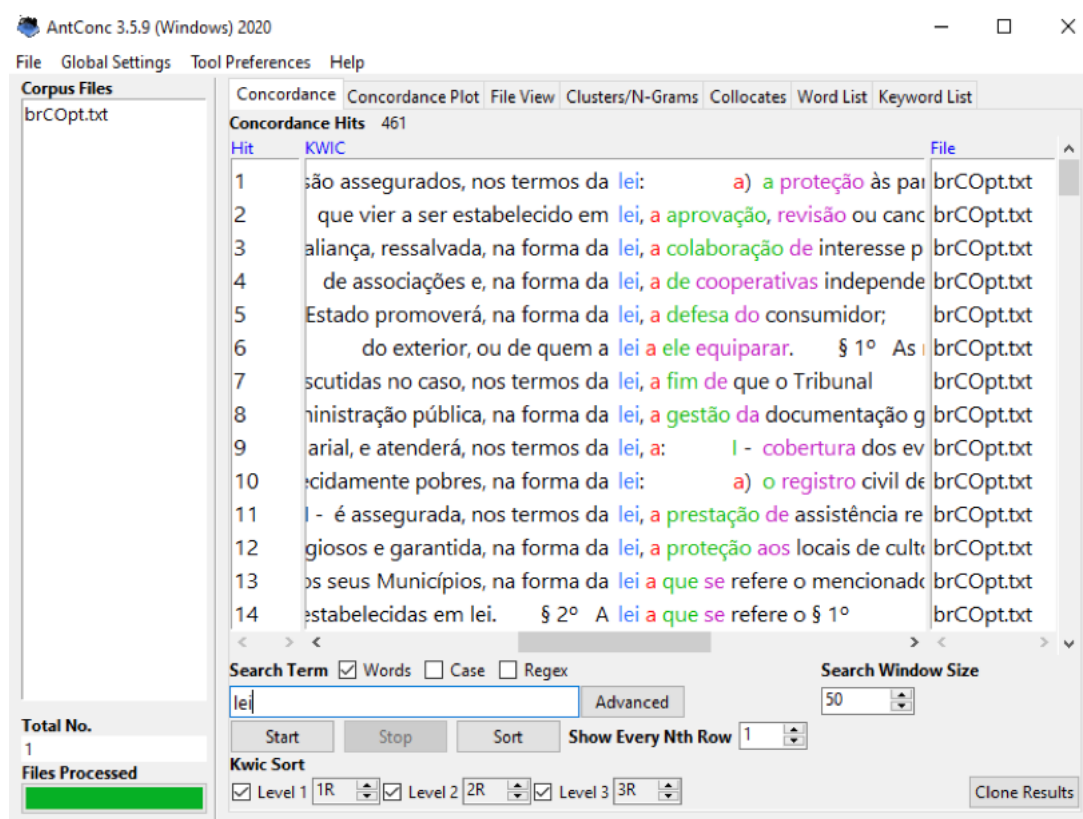


III. Concordance (concordância, contexto)

A lista gerada na aba *Concordance* detalha cada um dos itens contidos na frequência de uma palavra ou expressão, apresentando ainda os elementos que o cercam na frase em que ocorre a palavra ou grupo de palavras, ou seja, seu contexto. Como exemplo, selecionamos a palavra “**lei**” para analisar seu contexto. Podemos observar na Figura 7 que os grupos de palavras “**nos termos da lei**” e “**na forma da lei**” parecem ser recorrentes. Essas recorrências podem ser confirmadas na lista de n-gramas: 79 ocorrências para “**na forma da lei**” e 35 ocorrências de “**nos termos da lei**” na CF88. Separadamente, cada palavra (nos, termos, da e lei ou na, forma, da e lei) não tem valor terminológico, mas agrupadas nessa forma fixa – **nos termos da lei e na forma da lei** –, passam a caracterizar, não uma terminologia, mas um uso fraseológico comum do campo jurídico em um dado contexto. Em alguns tipos de RT, reconhecer fraseologismos desse tipo e também os que incluem ou se ligam às terminologias são bastante úteis pois caracterizam um “modo de dizer especializado”.

IMAGEM 6

Aba N-grams da ferramenta AntConc, com tamanho de 2 a 4



3.2. Exemplo de ficha terminológica

A ficha terminológica é um documento que, em um RT, corresponde ao conjunto ou registro individual com todas as informações que são coletadas sobre cada um dos termos ou expressões que correspondam a conceitos em um sistema de noções. É um documento muito útil, sobretudo porque ele se conectará a um dado texto ou *corpus*. A Ficha traz uma série de informação sobre um dado TERMO, SEU CONCEITOS, SUAS VARIABILIDADES e

seus registros conforme as fontes de apoio utilizadas no RT – tais como dicionários.

Apenas a título de breve exemplo, sugerimos o seguinte desenho para esse documento. Supomos uma ficha terminológica – em meio a um RT – que parte de textos legislativos com a nossa CF88. Vale comentar que hoje em dia já dispomos de diferentes ferramentas e plataformas computacionais especializadas na produção e armazenamento desse tipo de arquivo ou de “dossiê” sobre um dado termo. São ferramentas que lidam, em uma mesma base, com *corpora* e geram listagens diversas. Também nos ajudam a estabelecer um formulário eletrônico para a nossa ficha, conectando-o a diferentes tarefas de um RT. Um exemplo desse suporte amplo, para apoiar diferentes tipos de RTs, em diferentes tarefas, é o sistema e-Termos: <<https://www.etermos.cnptia.embrapa.br/index.php>>.

Mas, mesmo que não se conte com tais recursos mais sofisticados, partindo de um *corpus* documental criteriosamente selecionado, a ficha terminológica, independentemente de formato, é peça fundamental em um RT. A seguir, uma sugestão inicial que comporta, sempre, adaptações:

Termo	impacto ambiental
Língua	português
Corpus	Corpus de referência: CF88
Contexto de uso no <i>corpus</i>	[...] exigir, na forma da lei, para instalação de obra ou atividade potencialmente causadora de significativa degradação do meio ambiente, estudo prévio de impacto ambiental, a que se dará publicidade Fonte: (1)
Definição – conforme o <i>corpus</i> ou outras fontes de apoio	Qualquer alteração das propriedades físicas, químicas e biológicas do meio ambiente, causada por qualquer forma de matéria ou energia resultante das atividades humanas que, direta ou indiretamente, afetam: I – a saúde, a segurança e o bem-estar da população; II – as atividades sociais e econômicas; III – a biota; IV – as condições estéticas e sanitárias do meio ambiente; V – a qualidade dos recursos ambientais. Fonte: (2)
Ver também – termos e conceitos relacionados no corpus de referência	estudo de impacto ambiental; preservação ambiental
Fontes pesquisadas nesta ficha	(1) CF88 (contexto), (2) Resolução CONAMA 1/86 (definição) links (1) link (2)
Autor da ficha	FFE (iniciais do elaborador/terminólogo para identificar uma pessoa em um grupo de trabalho)
Data(s) da criação/ revisão desta ficha	07-03-2022 15-03 -2022

4. Perspectivas e desafios

Produzir um RT, mesmo havendo apoio computacional, nunca será uma tarefa trivial. Isso deve ter ficado evidente na parte introdutória deste texto. A ferramenta AntConc, como um recurso tradicional dos estudos de LC, mostra-nos o quanto determinadas contagens de palavras ao longo de um texto podem ser úteis para uma primeira abordagem a um conjunto documental. Mas há muito mais envolvido.

Ainda assim, elementos como as frequências de palavras, em listagens, nos revelam um olhar panorâmico especial. E, ao observarmos os grupos de palavras com três ou quatro elementos repetidos, vemos com mais nitidez, como se poderia “pescar” as terminologias e também as fraseologias ou as construções recorrentes do “lago” que um texto ou *corpus* representaria. Como candidatos a itens terminológicos recorrentes na CF88, podemos citar algumas duplas de palavras que correspondem aos conceitos de CRIMES INAFIANÇÁVEIS ou EXTRADIÇÃO DE ESTRANGEIROS, embora suas frequências sejam baixas e bem localizadas ao longo de todo o texto. Já a expressão “NA FORMA DA LEI” parece apontar para uma fraseologia típica desse tipo de documento e discurso.

Naturalmente, vale frisar, mostramos algumas contagens em separado de elementos que poderiam ser reunidos, como os plurais e os singulares, e as flexões de palavras. A ferramenta AntConc considera, por exemplo, que as palavras CRIME e CRIMES ou DE e DOS correspondem a dois elementos diferentes, sem considerar uma raiz ou base comum. Isso impactará na observação da TTR e de outros elementos sobre frequências de palavras e de expressões.

Esperamos, com este texto, que o nosso leitor se sinta animado a constituir o seu próprio *corpus* documental e a explorá-lo, tentativamente, com o AntConc. Uma sugestão interessante seria, por exemplo, comparar os vocabulários das diferentes Constituições dos países de língua portuguesa – todas são acessáveis no site do grupo TERMISUL. Qual delas teria a maior TTR? Quais seriam os grupos de quatro ou de três palavras mais repetidos ao longo das nossas constituições nacionais? Como um item como MULHER(ES) apareceria empregado ao longo desses documentos?

Questões como essas, e ainda muitas outras tantas, poderiam ser de muita utilidade quando se pretender iniciar um RT ou mesmo uma exploração linguística desinteressada em terminologias, seja com *corpora* legislativos ou com outros.

5. Agradecimentos

A produção deste texto contou com a participação de pessoas cujos estudos e pesquisas, realizados em universidades públicas do Brasil, são subvencionados por diferentes órgãos oficiais de fomento à pesquisa. Assim, registramos o nosso agradecimento ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), à CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), ao programa de Iniciação Científica e Tecnológica da FAPERGS (Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul) e ao Programa de Pós-Graduação em Letras da UFRGS. Fica também nosso obrigado especial à equipe do Grupo TERMISUL, que prossegue na senda do oferecimento público e gratuito de recursos e materiais diversos para o trabalho terminológico. Nossos agradecimentos a todo o grupo, na pessoa de sua coordenadora atual, a Profa. Dra. Cleci R. Bevilacqua. Por fim, agradecemos ao Departamento Cultural do Ministério das Relações Exteriores do Brasil,

Palácio Itamaraty e ao IILP pelo privilégio de poder participar no projeto Terminologias Científicas e Técnicas Comuns (TCTC) da Língua Portuguesa, cujas atividades de formação técnica, em 2021, ensejaram a escrita deste texto.

6. Referências Bibliográficas

ANTHONY, L. (2014): *AntConc* (Versão 3.4.3) [Software de computador]. Tóquio, Japão: Universidade de Waseda. Disponível em: <<http://www.laurenceanthony.net/>>. Acesso em: 7 mar. 2022.

BERBER SARDINHA, T. *Linguística de Corpus: histórico e problemática*. D.E.L.T.A., vol. 16, n. 2, p. 323-367, 2000. Disponível em: <<https://www.scielo.br/j/delta/a/vGknQkZQGsgYbrQfKmTZY4s/?format=pdf&lang=pt>>. Acesso em: 7 mar. 2022

BRASIL. Constituição (1988). *Constituição da República Federativa do Brasil*. Brasília, DF: Senado Federal, 1988. Disponível em: <http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm>. Acesso em: 7 mar. 2022.

BRASIL. Ministério da Mulher, da Família e dos Direitos Humanos. Lei nº 8.069, 13 de julho de 1990. Dispõe sobre o Estatuto da Criança e do Adolescente e dá outras providências. Diário Oficial da União, 1990. Disponível em: <<https://www.gov.br/mdh/pt-br/centrais-de-conteudo/crianca-e-adolescente/estatuto-da-crianca-e-do-adolescente-versao-2019.pdf>>. Acesso em: 7 mar. 2022.

CABRÉ, M. T. *A Terminologia, uma disciplina em evolução: passado, presente e alguns elementos de futuro*. Debate Terminológico, Porto Alegre, n. 1, p. 1-14, 2005. Disponível em: <<https://seer.ufrgs.br/riterm/article/view/21286/15349>>. Acesso em: 7 mar. 2022.

CONCEIÇÃO, M. C.; ZANOLA, M. T. (org.). *Terminologia e mediação linguística: métodos, práticas e atividades*. Faro, Portugal: Universidade do Algarve Editora, 1. ed., 2020. Disponível em: <<https://sapientia.ualg.pt/handle/10400.1/15043>>. Acesso em: 7 mar. 2022.

FINATTO, M. J. B. *Acessibilidade textual e terminológica: promovendo a tradução intralinguística*. Estudos Linguísticos (São Paulo. 1978), [S. l.], v. 49, n. 1, p. 72-96, 2020. DOI: 10.21165/el.v49i1.2775. Disponível em: <<https://revistas.gel.org.br/estudos-linguisticos/article/view/2775>>. Acesso em: 7 mar. 2022.

KRIEGER, M. G.; MACIEL, A. M. B.; ROCHA, J. C. C. R.; FINATTO, M. J. B.; BEVILACQUA, C. R. *Dicionário de direito ambiental: terminologia da legislação do meio ambiente*. 1. ed. Porto Alegre: Editora da UFRGS, 1998. v. 1. 511p.

KRIEGER, M. G.; FINATTO, M. J. B. *Introdução à Terminologia: teoria e prática*. São Paulo: Contexto, 2004.

MOTTA, E. *Sentenças Judiciais e Acessibilidade Textual e Terminológica*. Domínios de Linguagem, [S. l.], v. 15, n. 3, p. 761-813, 2021. DOI: 10.14393/DL47-v15n3a2021-6. Disponível em: <<https://seer.ufu.br/index.php/dominiosdelinguagem/article/view/52909>>. Acesso em: 7 mar. 2022.

Língua e diplomacia – Uma agenda para o diálogo em língua portuguesa

Margarita Correia

Sobre o autor

Universidade de Lisboa e CELGA-ILTEC. Presidente do Conselho Científico do IILP (2018-2022).

Resumo

Neste trabalho, são apresentados argumentos em favor da existência explícita de uma forma específica de diplomacia, a “diplomacia linguística”, e de formação específica nesta área para os diplomatas. O escopo da reflexão é Portugal, os países de língua portuguesa e especialmente a CPLP, organização internacional cujo fundamento é o facto de todos os países que a constituem terem o português como língua oficial. Chama-se, ainda, a atenção para a insustentável situação estatutária do Instituto Internacional da Língua Portuguesa (IILP), o órgão que, na CPLP, tem a seu cargo a gestão e difusão do português, por aquela situação ser o reflexo da incompetência ou do descaso dos Estados-Membros relativamente à língua que os (des)une. Por fim, ensaia-se uma definição do que poderia ser “diplomacia cultural e apresentam-se sugestões de ação, por ela enquadradas, para o espaço da Comunidade.

Palavras-chaves

Diplomacia, diplomacia linguística, língua portuguesa, CPLP, IILP

1. Um pouco da minha história pessoal e linguística – Percursos de quem (e/i)migra

Os meus pais, portugueses do distrito de Aveiro, emigraram para a Venezuela nos anos 1950, primeiro o meu pai, em 1955, e depois a minha mãe, em 1958. Fizeram-no por não ter encontrado no seu país as condições mínimas para construir a sua vida com o conforto a que aspiravam e acontece ainda hoje com tantos outros milhões de portugueses que, ao longo da história, a tal se viram obrigados¹.

Eu sou lusodescendente. Nasci na Venezuela em 1960. Não tenho a certeza de qual terá sido a minha língua materna, mas penso que terá sido um registo de português europeu com forte influência de castelhano². É preciso lembrar que, quanto menor for o índice de literacia dos indivíduos, maior é a sua permeabilidade a aculturação. Os portugueses residentes na Venezuela costumam chamar ao seu registo linguístico informal 'portunhol' e, como tal, partirei do princípio de que a minha língua materna foi esse portunhol. O meu portunhol era, ainda, muito influenciado pelo português brasileiro da minha madrinha, natural do Ceará, assim como pelas línguas e registos dos amigos que frequentavam a casa dos meus padrinhos: cubanos, espanhóis (incluindo bascos), italianos, franceses, libaneses, ...

Os meus pais até que gostariam de me ter posto a estudar em português (sempre foi sua intenção regressar a Portugal), mas era impossível na época, pois a comunidade portuguesa em Valência, ao contrário de outras comunidades, não se encontrava organizada, provavelmente devido aos baixos ou nulos índices de literacia da generalidade dos seus membros. De resto, até aos 22 anos tive apenas nacionalidade venezuelana, porque, apesar da distância hoje curta entre Valência e Caracas, o meu pai não teve possibilidade de se deslocar para me registar na Embaixada de Portugal e, na época, não havia Consulado em Valência, ao contrário do que acontece atualmente³.

Fui para o colégio em 1965, com cinco anos já feitos, e aí fui confrontada, creio, pela primeira vez, com castelhano formal. É preciso esclarecer que o pessoal do colégio era constituído por hispanófonos de Espanha, Venezuela, Cuba e talvez de outras proveniências. Lembro-me até de uma freira angolana, "la madre Fernanda", mas não recordo como era o seu registo linguístico. Entre as minhas colegas havia filhas de venezuelanos, espanhóis, italianos, alemães, franceses, libaneses, sírios, sobretudo católicos, mas lembro-me de pelo menos uma colega, a Irma, ser filha de judeus praticantes. Só com a minha entrada para a escola e consequente alfabetização em língua castelhana, acredito ter começado a expressar-me, pensar e sonhar em espanhol.

Como se pode verificar, até aos 10 anos de idade, vivi no meio de uma doce algarviada, um verdadeiro caldo de cores, culturas e diversidade. Só não sabia que, no país onde vivia, havia também línguas indígenas vivas, o que só foi reconhecido oficialmente muito mais

¹ Para uma panorâmica histórica da emigração portuguesa, vd. Serrão, J. 1970. Para uma visão mais recente da emigração portuguesa, vd. Gois, P.; Marques, J. C. 2018.

² Na Venezuela, a língua tem o nome oficial de "castelhano" – vd. nota 5.

³ O Consulado Geral de Portugal em Valência iniciou a sua atividade no dia 1 de agosto de 1988.

tarde⁴. E, apesar de tudo, sobrevivi! E considero-me muito privilegiada por ter tido esta experiência.

No início de 1971, deu-se o regresso antecipado dos meus pais a Portugal, devido à insegurança na Venezuela, que, como se vê, não é só de agora. Não vou alongar-me nas dificuldades de uma filha única, urbana, hispanofalante, vinda dos trópicos, que subitamente vai parar a uma aldeia do interior do distrito de Aveiro, para uma casa sem eletricidade, água corrente ou casa de banho (no sentido atual do termo), em um país frio, pobre, fechado, cinzento, triste e resignado. Falarei da experiência escolar e linguística.

Cheguei a Portugal a 14 de março de 1971, tendo estado nesse ano letivo a frequentar o “5.º grado” na Venezuela. Em Portugal não havia qualquer apoio estatal ou social ao regresso de emigrantes, como não havia, de resto, apoio para coisa alguma. Como era meio de ano letivo, a minha família não considerou a possibilidade de eu ingressar de imediato na escola, mas, por um golpe de sorte, a professora primária da minha prima, a Sra. D. Irene, falou com os meus pais, fez-lhes ver que seria um desperdício eu ficar fora da escola e ajudou-os nos trâmites para o meu ingresso. Sou-lhe eternamente grata!

Em julho desse ano, concluí a 4.ª classe, com a classificação de Bom. Note-se que, quando cheguei à escola portuguesa, não falava nem escrevia em português, embora compreendesse parte do que ouvia e conseguisse ler alguma coisa. A partir daí, o português, língua dominante – e única! – neste para mim novo país, passou a ser também a minha língua dominante, aquela em que, com mais frequência, me expresso, penso e sonho. O castelhano teve que ficar encerrado num qualquer recanto da minha memória e só muito mais tarde foi libertado, quando retomei contactos regulares com hispanofalantes.

Considero-me bilingue, pois, apesar das muitas interferências entre as duas línguas – inevitáveis, mesmo apesar de eu ser linguista e, por isso, conhecer a gramática explícita de ambas as línguas –, consigo ser proficiente em português e em castelhano quer ao nível da compreensão, quer ao nível da produção.

Tornei-me, sucessivamente, professora de português e linguista. Ao longo da minha carreira profissional de mais de 40 anos, tenho sempre pugnado pelo diálogo no seio dos países da Ibero-América e no dos países de língua portuguesa, e uns com os outros.

2. Os meus contactos com os diplomatas - Da perplexidade e do desconcerto

2.1. Considerações prévias

Nunca me ocorreu ser diplomata. Sempre considerei que a profissão estaria reservada a pessoas de classe social mais elevada do que aquela a que pertenço e, como tal, eu não teria acesso a ela, à partida; além disso, sou mulher. Acredito ainda hoje não estar longe da realidade relativamente à origem dos diplomatas, especialmente nos países que constituem

⁴ A Constituição da República Bolivariana da Venezuela, de 1999, no seu artigo 9.º estabelece o seguinte: “El idioma oficial es el castellano. Los idiomas indígenas también son de uso oficial para los pueblos indígenas y deben ser respetados en todo el territorio de la República, por constituir patrimonio cultural de la Nación y de la humanidad.”

a CPLP⁵. Porém, do breve relato aqui feito, do meu interesse por política linguística e da minha experiência nos anos de colaboração com o IILP (2010-2022), é fácil inferir o porquê de a questão diplomática não estar distante das minhas preocupações.

Até 2010, os meus contactos com diplomatas haviam sido escassos, limitados a contactos com representações diplomáticas de Portugal e da Venezuela, para tratar de assuntos pessoais ou familiares, e, ainda, a encontros fortuitos em eventos oficiais relacionados com a língua portuguesa. A partir de 2010, porém, iniciada a minha colaboração com o IILP, estes contactos passaram a ser profissionais e mais frequentes.

Antes de prosseguir, importa esclarecer que as minhas observações sobre a diplomacia e os diplomatas contidas neste texto não visam indivíduos particulares ou representantes de países particulares e que elas são elaboradas da perspectiva de uma pessoa bilingue e filha de emigrantes, profissional que tem por objeto de trabalho a linguística, especialmente a linguística portuguesa e, ainda, por uma cidadã que acredita que, apesar das origens que possam ter, os diplomatas são acima de tudo servidores do Estado (funcionários públicos, portanto) e que devem exercer a sua missão com profissionalismo e a isenção, tendo sempre em mente o bem comum. No entanto, o ponto de fuga desta perspectiva situa-se em Portugal, facto que se repercute no que adiante direi.

2.2. Dos diplomatas em geral

A primeira perplexidade de um comum mortal, quando contacta com diplomatas por razões profissionais, é constatar que estes parecem viver alheados do mundo real, do mundo de que fazem parte os Estados que servem e as pessoas que neles vivem ou deles são cidadãos. Em geral, os diplomatas, além de provirem de classes sociais altas, auferem salários e subsídios muito acima da média dos respectivos países⁶, têm proteção jurídica, social e até imunidade pelo exercício das suas funções, gravitam entre governantes, gente influente e poderosa e entre os seus pares, obedecem a protocolos muito rígidos. Não é difícil, portanto, compreender as razões para o alheamento referido. Porém, o facto de se compreender as putativas razões para esse alheamento não pode torná-lo aceitável nem desculpabilizá-lo, sobretudo tendo em conta que a condição básica dos diplomatas é serem servidores de Estado e o Estado sermos todos.

A par da atitude descrita, vem a linguagem, diferente, ininteligível para o comum dos mortais. Para entrar na carreira diplomática e singrar nela, é preciso ter, e dar provas disso, um vasto repertório de conhecimentos variados e aprofundados, em áreas como política, relações internacionais, jurisprudência, história, entre outras, o que pode levar a considerar a linguagem dos diplomatas como fruto do seu elevado índice de instrução e especialização.

⁵ Não foi tarefa fácil encontrar dados objetivos sobre a origem dos diplomatas em países de língua portuguesa. Relativamente ao estrato social de origem dos diplomatas brasileiros, encontrou-se informação aqui: <https://noticias.uol.com.br/colunas/democracia-e-diplomacia/2021/07/26/quem-representa-a-diplomacia-brasileira.htm>, que também refere a representatividade de género. Sobre esta questão, relativamente ao Itamaraty, vd. Amparo et al. 2021; em relação a Portugal, o texto de opinião de Augusto Santos Silva, ainda enquanto Ministro dos Negócios Estrangeiros de Portugal tem também interesse: <https://www.jn.pt/opiniao/convidados/a-igualdade-de-genero-na-diplomacia-portuguesa--13430194.html>;

⁶ De acordo com a sondagem “A opinião dos portugueses sobre a diplomacia portuguesa”, da Aximage, levada a cabo por entrevista telefónica nos dias 26 a 28 de outubro de 2018, 44,5% dos inquiridos considerou que os diplomatas portugueses são “demasiadamente bem pagos” (cf. Moita et al., coords., 2019, p. 169).

Além disso, eu bem sei que a linguagem hermética não é exclusiva dos diplomatas: políticos, agentes culturais, gente dos negócios e da alta finança, altos funcionários, intelectuais e outros profissionais recorrem a linguagem hermética para se afirmarem socialmente e demonstrarem a sua pertença de classe. Há quem o faça especificamente para entender os outros, mas não ser entendido por eles. Contudo, um falante efetivamente eficiente de uma língua, como se espera que sejam os diplomatas, é, entre outros fatores, aquele que domina perfeitamente diferentes registos linguísticos, dos mais elaborados aos neutros e mesmo aos mais simples, e sabe escolher o registo linguístico adequado à situação, ao interlocutor, ao conteúdo e aos objetivos da comunicação. Falar bem é, assim o entendo, uma questão de adequação discursiva.

Os diplomatas enfermam também de uma confusão, generalizada socialmente, entre o que são os domínios específicos da cultura, da literatura e da linguística⁷. Para uma melhor compreensão do argumento, importa definir o modo como os conceitos de cultura, literatura e linguística estão a ser usados neste trabalho. 'Cultura' é conceito mais abrangente e aqui entendida como "sistema complexo de códigos e padrões partilhados por uma sociedade ou um grupo social e que se manifesta nas normas, crenças, valores, criações e instituições que fazem parte da vida individual e coletiva dessa sociedade ou grupo"⁸. A língua é não só parte integrante, mas, pela sua natureza, aspeto determinante da cultura de uma sociedade. O mesmo dicionário apresenta duas aceções que nos interessam neste ponto para a palavra 'literatura': "2. conjunto de produções literárias de um país ou de uma época; (...) 4. disciplina que tem por objeto de estudo os estudos literários [sic]"⁹. Por fim, ainda de acordo com a mesma fonte, 'linguística' é a "ciência que tem por objeto de estudo a linguagem humana, desde o plano da língua [enquanto sistema interiorizado de regras e vocábulos] até ao plano do discurso [enquanto realização concreta e irrepetível da linguagem verbal em qualquer registo]"¹⁰. É inegável que a língua é objeto de trabalho da literatura e da linguística, mas uma e outra têm conceções, abordagens e metodologias de trabalho diferentes sobre a sua matéria-prima: a literatura é uma forma de arte e tem sobre a língua uma conceção fundamentalmente estética; a linguística, por seu turno, é uma ciência e tem sobre a língua uma conceção teórica e aplicada. Como tal, a literatura está mais próxima da cultura (assim como outras formas de arte) e a linguística mais próxima das necessidades e da prática social. Literatura e linguística não são antagónicas, nem conflituantes e nenhuma é mais importante que a outra: ambas são complementares e importantes, mas diferentes.

⁷ Desta confusão é exemplo paradigmático o Programa do XXIII Governo Constitucional da República Portuguesa (2022-2026), que inclui, no apartado "I.VI. Cultura", o subtítulo "Internacionalizar as artes e a língua portuguesa" (p. 177-178). No restante documento, a expressão "língua portuguesa" não volta a surgir em nenhum título, tendo mais 11 ocorrências dispersas. Disponível em <https://www.portugal.gov.pt/pt/gc23/governo/programa-do-governo>. Ainda a propósito desta confusão de conceitos, vd. e.g. <https://www.diplomaciabusiness.com/o-que-e-diplomacia-cultural-paises-ganham-influencia-atraves-de-musica-e-cinema/>

⁸ Porto Editora – cultura no Dicionário infopédia da Língua Portuguesa [em linha]. Porto: Porto Editora. [consult. 2022-08-16 13:06:37]. Disponível em <https://www.infopedia.pt/dicionarios/lingua-portuguesa/cultura>.

⁹ Porto Editora – literatura no Dicionário infopédia da Língua Portuguesa [em linha]. Porto: Porto Editora. [consult. 2022-08-16 13:13:54]. Disponível em <https://www.infopedia.pt/dicionarios/lingua-portuguesa/literatura>.

¹⁰ Porto Editora – linguística no Dicionário infopédia da Língua Portuguesa [em linha]. Porto: Porto Editora. [consult. 2022-08-16 13:18:37]. Disponível em <https://www.infopedia.pt/dicionarios/lingua-portuguesa/linguistica>.

Apesar dos profundos conhecimentos de história nacional e internacional e de relações político-culturais entre países que os diplomatas seguramente têm, é muito surpreendente o grande desconhecimento que revelam relativamente às línguas que falam ou com que convivem por força da sua atividade, às suas sócio-histórias, à forma como configuram o tecido social e cultural, ao tipo de relações que estabelecem com outras línguas e à sua relevância estruturadora do pensamento e identidade. É também preocupante o desconhecimento de conceitos básicos sobre a língua, como os de norma e variação, variedade, registo linguístico, adequação discursiva, recursos linguísticos¹¹, política linguística, enfim, a ausência de um olhar crítico sobre os usos linguísticos das sociedades que as falam. É certo que a linguística é uma ciência relativamente nova, que, na sua moderna aceção, nasceu com o século XX, mas, a tecnologia digital, os meios de comunicação e transporte atuais, o turismo, as organizações internacionais, a mediatização são seus contemporâneos e, ainda assim, são conhecidos e valorizados. É angustiante, em suma, o desconhecimento que os diplomatas, e os Governos que determinam a sua ação, tantas vezes exibem relativamente à existência de uma ciência chamada 'linguística', à língua e à sua importância estratégica, política, social, cultural, i.e., ao seu papel em todas as facetas da atividade humana¹².

2.3. Dos diplomatas da CPLP e da sua relação com a língua portuguesa e com o IILP

Se passarmos ao contexto mais restrito da CPLP, comunidade cimentada na partilha de uma língua comum e agregadora, chega a ser chocante o descaso a que é votada a língua portuguesa, apesar de ser alicerce estruturador da própria Comunidade, em detrimento dos demais pilares da sua atuação, a saber, o político-diplomático e o da cooperação¹³, i.e., as áreas que envolvem diretamente dinheiro e poder. Importa destacar que as críticas à CPLP, que são muitas, pertinentes e expressas, deverão ser endereçadas, em primeira instância, aos governos e administrações públicas dos seus Estados-membros. Os diplomatas, contudo, não podem assumir-se como meros executores acríticos das políticas dos seus governantes, mas poderão, em virtude da sua condição de elite e dos horizontes amplos que lhes é dado enxergar, ter um relevante papel na reflexão e no aconselhamento dos Estados que servem.

¹¹ Posto que esta é a única questão que me foi colocada após a apresentação oral desta reflexão, aqui fica uma possível definição sucinta do conceito de 'recursos linguísticos': instrumentos disponíveis aos falantes da língua como auxiliares no uso adequado ou no ensino da mesma. São exemplos de recursos linguísticos as gramáticas, os dicionários, os vocabulários ortográficos, as terminologias científicas e técnicas.

¹² Para este desconhecimento, muito contribuem também os próprios linguistas, que não têm, pelo menos em Portugal, intervenção social e política significativas.

¹³ De acordo com os Estatutos da CPLP (revisão de 2007), "Artigo 4º (Objetivos) – 1. São objetivos gerais da CPLP: a) A concertação político-diplomática entre os seus membros em matéria de relações internacionais, nomeadamente para o reforço da sua presença nos fora internacionais; b) A cooperação em todos os domínios, inclusive os da educação, saúde, ciência e tecnologia, defesa, oceanos e assuntos do mar, agricultura, segurança alimentar, administração pública, comunicações, justiça, segurança pública, economia, comércio, cultura, desporto e comunicação social; c) A promoção e difusão da Língua Portuguesa, designadamente através do Instituto Internacional de Língua Portuguesa."

Um dos termos que me ocorre para descrever a atuação da CPLP é 'umbiguismo', i.e., a característica da pessoa que se concentra apenas em si própria. Este umbiguismo comunitário manifesta-se a dois níveis. 1) Ao nível de cada um dos Estados-membros, que têm muitas vezes dificuldade em apreender que comunidade implica coexistência, cooperação, respeito mútuo e valorização do bem comum, a par do bem nacional: será muitas vezes difícil conciliar o bem comunitário com o bem nacional, mas criar uma comunidade de países com base numa língua comum não é para fracos¹⁴ e os diplomatas da CPLP devem ter sempre presente a sua responsabilidade maior. 2) Mas o umbiguismo é também apanágio da Comunidade no seu todo, que parece empenhada em descobrir a roda ou a pólvora, não se detendo para estudar aquilo que outras organizações internacionais, outros países e outras sociedades têm feito em prol das suas línguas. Impressiona, e.g., que uma comunidade constituída por países multilingues aparente ter tão pouco interesse na promoção do multilinguismo nos seus países, a par do crescimento do português, aspeto que constitui, desde há décadas, um dos desígnios mais importantes das Nações Unidas, particularmente através da UNESCO.

A CPLP não é uma entidade benquista nem respeitada pelas sociedades dos países que representa. É certo que a afirmação anterior se baseia apenas no que se lê e escuta na comunicação social e também nas opiniões das pessoas com quem vou falando, mas não me surpreenderia que uma sondagem sobre o assunto, levada a cabo nos países da CPLP (porque não fazê-la?), revelasse uma realidade semelhante. Para o comum dos mortais, não são claros os objetivos a alcançar pela CPLP, as metas definidas, os benefícios a conquistar, o rumo a seguir, e, às vezes, parece que nem a própria CPLP tem sempre estes dados muito presentes.

Bem sei que há alguma dose de injustiça nesta apreciação, porque o trabalho que se faz na instituição não tem visibilidade pública e antes decorre nos bastidores político-diplomáticos, mas também é verdade que em mais de 25 anos de existência poderão contar-se pelos dedos da mão os êxitos relevantes para os países que a constituem¹⁵. Além disso, quando se tem a oportunidade de conhecer o funcionamento da instituição, não é possível deixar de experimentar algum desencanto e desconforto relativamente à sua eficiência e eficácia. Posta esta disparidade aparente entre as intenções explícitas e as implícitas da CPLP, assim como o distanciamento que a instituição parece manifestar relativamente às sociedades que serve, não é também de estranhar a comprovada ausência de motivação da maioria dos académicos dos diferentes países da CPLP para participar em projetos que, além de não trazerem benefícios pecuniários e dependerem de vontades políticas efémeras, são em geral pouco interessantes cientificamente, não granjeiam reconhecimento no mundo

¹⁴ De acordo com Moita et al., coords., (2019: 26-27), um dos "momentos de especial significado e impacto que, no seu conjunto, estimularam e consolidaram uma evolução serena, mas acentuada e positiva da diplomacia portuguesa" (p. 26) foi a criação da CLP em 2019, que "veio introduzir um desafio adicional. Desde logo pelos propósitos definidos para a novel organização, mas também pela exigência de, sob uma ótica de interesses partilhados, trabalhar multilateralmente com os novos Estados africanos independentes e de ter o Brasil como relevante parceiro nesse propósito." (p. 27).

¹⁵ Note-se, e.g., que só ao fim de 25 anos foram estabelecidos os princípios básicos da mobilidade dos cidadãos, mas que estes estão longe de poder ser aplicados em todo o espaço comunitário. "O «Acordo sobre a Mobilidade entre os Estados-Membros da CPLP» [que é um Acordo-quadro que estabelece a base legal sobre a qual se construirá uma maior mobilidade e circulação no espaço da CPLP] foi aprovado na XXVI Reunião do Conselho de Ministros da Comunidade dos Países de Língua Portuguesa (CPLP), decorrida, em Luanda, Angola, no dia 16 de julho de 2021." ([https://secretariadoexecutivo.cplp.org/comunicacao/noticias/noticia-detallhe/?news=6648](https://secretariadoexecutivo.cplp.org/comunicacao/noticias/noticia-detalhe/?news=6648)). Note-se também que continua a não existir regulação que possibilite a livre circulação de bens culturais e de conhecimento científico e tecnológico na CPLP.

competitivo dos índices bibliométricos e das avaliações do desempenho e nem sequer no seio das sociedades em que vivem.

A CPLP é o espelho dos Estados que a constituem e da eficácia governativa dos seus governos.

Antes de a CPLP ser criada, em 1996, foi concebido o IILP, em 1989, e este foi, sem dúvida o embrião do que viria a ser aquela¹⁶. Tendo em conta esta inegável realidade, torna-se incompreensível o enorme descaso a que é votado o IILP, os seus órgãos de gestão e funcionamento, assim como os seus colaboradores¹⁷, quer pela CPLP, quer pelos Estados-membros¹⁸. Estes fazem-se representar no IILP através das suas Comissões Nacionais, que, por seu turno constituem aquilo a que, pelos Estatutos aprovados em 2005, se chama o “Conselho Científico”. Contrariamente ao que seria de esperar, as Comissões Nacionais são raramente constituídas por cientistas ou académicos, especialistas em linguística.

No IILP assistimos a muito do já foi referido acima a propósito do funcionamento da CPLP, mas, dadas as especificidades do instituto – a sua proverbial falta de recursos, a sua condição de “patinho feio” da Comunidade (porque não é rentável, não gera poder nem dinheiro, porque muitos não entendem para que serve, nem as suas potencialidades) -, são mais evidentes as rivalidades entre países, o paternalismo dos mais ricos em relação aos mais pobres, a dificuldade de afirmação dos mais pequenos relativamente aos grandes ou mais poderosos¹⁹. Este cenário é tanto mais triste e lamentável quanto são escassos e obtidos com esforço os poucos recursos de que o IILP dispõe. Este tipo de atuação gera desconfiança generalizada das sociedades em relação ao potencial de desenvolvimento do Instituto, à pertinência, credibilidade e transparência da CPLP, aos políticos e à diplomacia em geral e, por extensão, em relação às virtudes dos regimes democráticos em que, desejavelmente, vivemos.

¹⁶ Cf. Processo Histórico – CPLP em <https://www.cplp.org/id-2752.aspx>

¹⁷ Refira-se que os funcionários do IILP não se encontram enquadrados na orgânica da CPLP, auferindo, por isso, salários e subsídios mais baixos, e conhecendo situações laborais bastante mais desfavoráveis e precárias.

¹⁸ O IILP tem duas versões distintas de Estatutos, ambas aprovadas em Conselho Científico e em Conselhos de Ministros da Comunidade dos Países de Língua Portuguesa (CPLP): a de 1998, alterada em 2005 (conhecidos como “estatutos de 2005”) e a de 2010. O Brasil e Portugal ratificaram os “estatutos de 2005” (cf. Decreto Legislativo N.º 288, de 2008, Brasil; Resolução da Assembleia da República, Diário da Assembleia da República – II Série A - Número: 16851, de 29 de julho de 2008, Portugal). Apenas Portugal ratificou todas as versões que os estatutos conheceram até hoje – Cf. Decreto 19/99 (Diário da República, 1.ª série – N.º 34 – 20 de agosto de 1999), que aprova os Estatutos Revistos do Instituto Internacional da Língua Portuguesa, adotados na Praia, Cabo Verde, aos 17 de julho de 1998; Decreto 12/2012 (Diário da República — N.º 115 — 15 de junho de 2012), que ratifica o Estatutos aprovados em 2010. Os “estatutos de 2005” resultam de alterações aos Estatutos originais do Instituto Internacional de Língua Portuguesa (IILP), aprovadas na X Reunião Ordinária do Conselho de Ministros da CPLP, realizada em Luanda, nos dias 19 e 20 de julho de 2005, de modo a articulá-los com os Estatutos da CPLP. A existência de Comissões Nacionais e Conselho Científico, referida no texto principal, está contida nos estatutos de 2005, que têm norteado até hoje a atividade do Instituto.

Por outro lado, há Estados-membros que nunca regularizaram os seus compromissos financeiros para com o IILP; há também aqueles nunca nomearam as suas Comissões Nacionais, ou que as mantêm há anos numa situação de indefinição. Por fim, há Estados-membros que nunca participam nos projetos do IILP, não nomeando equipas nacionais para esses projetos ou não lhes dando condições para o desenvolvimento da sua atividade. Todas estas situações são prova clara do descaso a que o IILP é votado por muitos Estados-membros da CPLP.

¹⁹ Em situações de cooperação entre o IILP e membros da sociedade civil, quando há uma mudança de estratégia de um ou mais países ou se alteram as relações de força entre eles, é possível até assistir a cenas pouco edificantes de acareação, despropositada e deselegante, das credenciais e competências dos linguistas que colaboram com o IILP e do trabalho por eles desenvolvido, como, infelizmente, já aconteceu.

A errática situação do IILP é tanto menos explicável e aceitável quanto os Estatutos aprovados em 2010, mas não em vigor por falta de ratificação pelos Estados-Membros, são exímios na forma como definem as atribuições do IILP, no Artigo 4.º:

2 — São atribuições gerais do IILP:

- a) Elaborar e executar planos estratégicos e consequentes programas e projetos conducentes à internacionalização efetiva da Língua Portuguesa, nomeadamente junto de organizações internacionais e regionais;
- b) Promover, propor e apoiar a execução de projetos culturais, numa perspetiva integrada e de interculturalidade, de difusão da Língua Portuguesa;
- c) Fomentar a promoção do conhecimento da Língua Portuguesa e das Culturas da CPLP, nomeadamente, através do estabelecimento de redes de investigação e de intercâmbio de experiências;
- d) Promover e acompanhar o desenvolvimento de instrumentos de normalização linguística comum;
- e) Propor e ou avaliar projetos multilaterais de apoio ao desenvolvimento e otimização das competências em Língua Portuguesa em Estados membros da CPLP;
- f) Apoiar a adequada articulação entre a Língua Portuguesa e as demais línguas nacionais;
- g) Assegurar a representação da CPLP em fora internacionais sobre multilinguismo e multiculturalismo, por delegação do Secretário Executivo da CPLP;
- h) Colaborar com instituições dos Estados membros e de Estados terceiros no desenvolvimento de programas e projetos relevantes para a Língua Portuguesa.

3. Tópicos para a construção de uma agenda para o diálogo em língua portuguesa e – Ideias soltas, para descartar ou não

3.1. Será que existe “diplomacia linguística”?

Uma das questões que me coloco com frequência é a de saber se existe e é usado o termo ‘diplomacia linguística’, associado a um conceito estabilizado, enquanto complementar, e.g., de ‘diplomacia cultural’, amplamente usado.

Da leitura de Moita et al. coords., 2019, verifica-se que a diplomacia portuguesa se desenvolve em torno de três eixos principais (diplomacia política, diplomacia económica e diplomacia cultural²⁰) e que a expressão “diplomacia linguística” não ocorre nenhuma vez.

Inevitavelmente, fiz uma busca no Google²¹, usando a expressão de pesquisa “diplomacia linguística”.pt e obtive os seguintes resultados: 35 ocorrências, com apresentação de 23 entradas, mas nem todas aproveitáveis. As 10 ocorrências e respetivos contextos retidos aparecem listados, de A. a H., no anexo a este trabalho. A sua análise permite-nos inferir que:

²⁰ Sobre a dificuldade de definir ‘diplomacia cultural’, cf. e.g. Španjević, M. 2014.

²¹ 17 horas e 55 minutos, do dia 16 de agosto de 2022.

1. diplomacia linguística é associada muitas vezes a diplomacia cultural, mas são coisas diferentes (F., H., I, K.);
2. se fala mais do assunto no Brasil do que em Portugal, ou pelo menos há mais trabalhos académicos que a referem naquele país (B., C., H.);
3. a diplomacia linguística aparece associada a entidades de promoção da língua no estrangeiro (F., J., K.);
4. para algumas ONG portuguesas, diplomacia linguística, associada ao Camões I.C.L., opõe-se a cooperação e é muito menos necessária e importante (F., K.)²²;
5. é confundida com cortesia linguística (E., G.);
6. pode ser entendida como a capacidade de usar várias línguas no mesmo discurso diplomático (A.);
7. diplomacia linguística deve estar associada a cooperação técnica e económica (C.);
8. pode sofrer incómodos e problemas quando a mesma língua é falada em vários países independentes (D.);
9. para um diplomata, a diplomacia linguística está associada à 'diplomacia migratória' (I.);
10. diplomacia linguística está associada a ensino de português no estrangeiro e a leitorados (H.).

O conceito parece não estar claramente definido, mesmo entre diplomatas, mas parece estar associado ao de diplomacia cultural e difusão do português no exterior, particularmente através do ensino. Não pode deixar de surpreender que, em nenhum excerto, a diplomacia linguística tenha sido associada, e.g., à língua portuguesa como língua de ciência ou ao seu uso no seio de organizações internacionais, embora estes temas sejam recorrentes nos diferentes Planos de Ação aprovados pelos Chefes de Estado e de Governo da CPLP²³.

Não tendo formação em política externa e em diplomacia, é na qualidade de linguista que tem desempenhado funções nesta área há 13 anos, que proponho que o conceito de diplomacia linguística seja alvo de estudo e aprofundamento, e que abranja outras perspetivas e áreas de atuação, além do ensino de português no estrangeiro.

No âmbito deste trabalho, proponho que diplomacia linguística seja provisoriamente entendida como: conjunto de estratégias, iniciativas e ações que visem promover o desenvolvimento, a difusão e o uso da língua, a nível internacional, nas mais distintas situações

²² As referências a esta oposição entre língua e cooperação resultam do facto de, em 2012, o Governo português ter decidido fundir o até então Instituto Camões, dedicado à difusão internacional da língua portuguesa, e o Instituto Português de Apoio à Cooperação (IPAD), exclusivamente ligado à cooperação e desenvolvimento. Esta fusão gerou forte controvérsia entre as ONG portuguesas. Dela resultou o "Camões, Instituto da Cooperação e Língua" ou "Camões I.C.L.". Em 2017, o Decreto Regulamentar n.º 3/ de 2017 (cf. Diário da República, 1.ª série — N.º 83 — 28 de abril de 2017) estabelece o regime jurídico do estatuto da empresa promotora da língua portuguesa Instituto, que passa a designar-se "Camões, I.P." e à Agência para o Investimento e Comércio Externo de Portugal, E. P. E.

²³ 1) Plano de Ação de Brasília para a Promoção, a Difusão e a Projeção da Língua Portuguesa, documento anexo à resolução da VI Reunião Extraordinária do Conselho de Ministros da CPLP, realizada em Brasília, em 31 de março de 2010. 2) Plano de Ação de Lisboa, aprovado XII Reunião Extraordinária do Conselho de Ministros da CPLP (Maputo, 20 de fevereiro de 2014). 3) Plano de Ação de Díli, adotado pela XII conferência de Chefes de Estado e de Governo da CPLP (Santa Maria, ilha do Sal, 17 e 18 de julho de 2018,). 4) Plano de Ação da Praia, adotado XIII Conferência de Chefes de Estado e de Governo da CPLP (Luanda, 16 de julho de 2021).

de uso (e.g. como língua materna ou adicional²⁴, de ensino e aprendizagem, de difusão de conteúdos, de negócios, de desenvolvimento cultural, científico e social).

3.2. Que objetivos e funções pode desempenhar a diplomacia linguística da língua portuguesa?

Para organização do raciocínio, partamos do princípio de que a atividade de um diplomata português se pode desenvolver, entre outros, nos seguintes contextos:

- 1) enquanto agente acreditado junto de Estados onde o português não é língua;
- 2) enquanto agente acreditado junto de organizações internacionais onde o português não é língua de trabalho;
- 3) enquanto agente acreditado em países com forte presença de comunidades portuguesas e de lusodescendentes;
- 4) enquanto agente acreditado junto de Estados da CPLP;
- 5) enquanto agente acreditado junto de ou em exercício de funções na própria CPLP.

Ainda que, em todos os contextos enunciados, a diplomacia linguística seja absolutamente imprescindível, os tópicos seguintes dizem respeito sobretudo aos contextos 4) e 5), acima. É importante, elencar as especificidades da situação geopolítica-linguística que a língua portuguesa conhece atualmente:

- a língua portuguesa é língua oficial de nove países, distribuídos por três continentes;
- apesar do estatuto comum, a língua portuguesa conhece índices de difusão, descrição, codificação e uso diferentes nos nove países;
- é uma língua pluricêntrica, com duas normas estabelecidas, como tal dominantes, e algumas normas nacionais emergentes;
- ela constitui a génese da Comunidade de Países de Língua Portuguesa (CPLP);
- a CPLP integra o Instituto Internacional de Língua Portuguesa (IILP), a quem foi cometida a obrigação de promover a língua portuguesa no espaço da CPLP e no exterior²⁵;
- desde 2010, os Estados-membros da CPLP optaram por discutir em conjunto o lugar da língua portuguesa no sistema mundial e elaboraram quatro planos de ação²⁶;

²⁴ 'Additional' applies to all, except, of course, the first language learned. An additional language, moreover, may not be foreign since many people in their country may ordinarily speak it. The term 'foreign' can, moreover, suggest strange, exotic or, perhaps, alien—all undesirable connotations. Our choice of the term 'additional' underscores our belief that additional languages are not necessarily inferior nor superior nor a replacement for a student's first language." – Judd et al. (2001: 6).

²⁵ De acordo com os Estatutos da CPLP (revisão de 2007): "Artigo 9º (Instituto Internacional de Língua Portuguesa) O Instituto Internacional de Língua Portuguesa (IILP) é a Instituição da CPLP, dotada de Estatutos próprios, que tem como objetivos a planificação e execução de programas de promoção, defesa, enriquecimento e difusão da Língua Portuguesa como veículo de cultura, educação, informação e acesso ao conhecimento científico, tecnológico e de utilização em fóruns internacionais." Este artigo reproduz *ipsis verbis* a alínea 2, do art.º 1.º dos estatutos do IILP (2005). Já de acordo com os Estatutos aprovados em 2010: "Artigo 4.º (Missão e atribuições) – 1 — O IILP tem por missão a promoção, a valorização e a difusão da Língua Portuguesa como: i) Língua de expressão de diferentes culturas; ii) Língua de comunicação global e de utilização oficial em fóruns internacionais; iii) Língua de educação, formação e informação; iv) Língua de conhecimento científico e tecnológico; v) Língua de negócios e de desenvolvimento socioeconómico."

- desde 2014, os Estados-membros reconheceram a importância dos recursos linguísticos comuns para o português, ao assumir o Vocabulário Ortográfico Comum da Língua Portuguesa (VOC) como património da CPLP²⁷;

- desde 2016, todos os Estados-membros assumiram o pluricentrismo como princípio orientador da sua política linguística externa, ao escolhê-lo como primeiro eixo temático da Conferência de Díli²⁸.

Deste modo, acreditamos que são objetivos e funções da diplomacia linguística da língua portuguesa, no âmbito da CPLP, entre outros, os seguintes:

- servir a CPLP e os seus Estados-membros;
- promover a coesão, o diálogo, o entendimento e o respeito entre os Estados-membros;
- promover o diálogo entre a CPLP e outras organizações internacionais;
- contribuir para o desenvolvimento e prestígio da CPLP na comunidade internacional e junto das sociedades que a constituem;
- contribuir para diminuir o fosso entre países pobres e países ricos dentro da CPLP;
- nomeadamente, contribuir para que todos os países da CPLP atinjam níveis de literacia e taxas de escolarização próprias de países desenvolvidos: em língua materna se diferente do português, mas assegurando sempre o acesso à literacia na(s) língua(s) oficial(is), i.e., em português;
- promover o respeito pelos direitos humanos, a cidadania e por princípios democráticos no espaço da CPLP.

3.3. Que ações podem ser desenvolvidas para fomentar a política linguística da língua portuguesa?

As propostas abaixo enunciadas visam não apenas o mundo da diplomacia e os diplomatas, mas pretendem constituir uma chamada de atenção aos governos e administrações de todos os Estados-membros da CPLP, de quem os diplomatas dependem, devendo por aqueles ser consideradas.

Para fomentar uma diplomacia linguística estruturada e eficiente no âmbito dos países de língua portuguesa, será necessário:

- reconhecer que a língua portuguesa é instrumento fundamental na construção da identidade e cidadania e no desenvolvimento económico e social dos Estados-membros e dos seus cidadãos;
- reconhecer que a língua portuguesa tem, sim, também valor económico e que este representa muito mais do que seu peso no PIB dos Estados-membros;

²⁶ Cf. nota 24.

²⁷ Cf. a Declaração de Díli, emanada da X Conferência de Chefes de Estado e de Governo da Comunidade de Países de Língua Portuguesa, de 23 de julho de 2014. Acessível aqui: <https://www.cplp.org/id-4447.aspx?Action=1&NewsId=3463&M=NewsV2&PID=10872>

²⁸ Cf. a apresentação da III Conferência Internacional sobre o Futuro do Português no Sistema Mundial, Díli, 15 a 17 de junho de 2016, cujo primeiro eixo se intitulou “Português, Língua Pluricêntrica do Século XXI” (<https://www.cplp.org/id-4447.aspx?Action=1&NewsId=4285¤tPage=7&M=NewsV2&PID=10872>).

- reforçar a pluricentricidade da língua portuguesa, adotando medidas de gestão partilhada, de promoção do respeito por todas as variedades nacionais, de apoio à afirmação e desenvolvimento das variedades nacionais emergentes;
- entender o que é o pluricentrismo linguístico como programa de atuação política e aquilo que ele implica;
- contribuir para diminuir o fosso entre os países mais pobres e os países mais ricos da CPLP, promovendo o desenvolvimento da língua em todos e cada um, e acesso a recursos linguísticos de qualidade;
- conhecer modelos de gestão de outras línguas, nomeadamente pluricêntricas, desenvolvidos por órgãos nacionais e internacionais, nomeadamente para deles extrair possíveis contributos para a gestão partilhada da língua portuguesa;
- entender verdadeiramente os problemas de natureza linguística que os cidadãos da CPLP enfrentam, fruto da ausência de definição de políticas linguísticas claras ou da sua conceção e aplicação incompetentes;
- procurar resolver efetivamente esses problemas.

O desenvolvimento das estratégias e ações elencadas acima implicam a existência de diplomatas especialmente aptos à sua execução, com especialização em diplomacia linguística e que:

- conheçam a história e o enquadramento das instituições que servem, nomeadamente da CPLP,
- conheçam a sócio-história do português e das línguas nacionais dos Estados-membros da CPLP,
- possuam noções de linguística aplicada²⁹, particularmente conhecimentos básicos de sociolinguística³⁰ e de política linguística³¹,
- consigam entender e apreciar a pluricentricidade do português,
- tenham um olhar crítico sobre a realidade linguística,
- ultrapassem preconceitos e generalizações apressadas relativamente aos outros e às línguas e que falam,
- respeitem as especificidade e competências das demais entidades e pessoas envolvidas na diplomacia linguística.

Haveria interesse em que esta especialização fosse apoiada ou até promovida pela própria CPLP, no âmbito das suas atividades de formação avançada de diplomatas, com apoio de instituições idóneas dos Estados-membros. Outra possibilidade seria o desenvolvimento de estudos de pós-graduados em Estados-membros que o desejassem, assegurando o envolvimento de massa crítica relevante para tal. Por fim, seria possível, em modelo menos estruturado, promover ciclos de palestras ou seminários direcionados aos diplomatas.

²⁹ Sucintamente, pode dizer-se que a linguística aplicada procura resolver problemas sociais envolvendo a língua, ou pelo menos contribuir para a sua resolução.

³⁰ A sociolinguística estuda a relação entre a língua e a sociedade, descrevendo o efeito de aspetos da sociedade, incluindo normas, linguísticas e culturais, expectativas e contexto, no modo como a língua é usada, assim como os efeitos do uso da língua na sociedade.

³¹ Política linguística pode ser sucintamente definida como o estudo da relação entre a(s) língua(s) e o poder.

Sendo a língua portuguesa o cimento agregador da CPLP e tendo sido o IILP o embrião da Comunidade, importa reconhecer e respeitar a contribuição inestimável do Instituto para o desenvolvimento e a difusão da língua portuguesa na CPLP e no plano internacional, nomeadamente:

- dotando-o do quadro estatutário adequado, imprescindível ao seu funcionamento pleno em legalidade, atualizado, aprovado por todos Estados-membros da CPLP e ratificado por cada um deles;
- integrando o IILP no quadro orgânico da CPLP;
- oferecendo aos funcionários do IILP condições de trabalho e remuneração semelhantes às dos colegas ao serviço da CPLP³²;
- atribuindo ao IILP recursos humanos e financeiros, regulares e suficientes para o seu funcionamento,
- financiando condignamente o funcionamento do IILP, de modo que este deixe de estar condicionado à receção de iniciativas assistencialistas e contribuições voluntárias dos Estados-membros, para poder sobreviver,
- reforçando a representatividade dos Estados-membros no IILP, investindo em representações com competências e mandato adequados e garantindo que cada um dos Estados apoia e sustenta o funcionamento das representações nacionais e a sua contribuição para o desenvolvimento dos projetos do IILP;
- apoiando o estabelecimento de parcerias estratégicas entre o IILP e instituições da sociedade civil, nomeadamente universidades e centros de investigação que se dedicam a investigação e desenvolvimento na área da linguística;
- dedicando ao IILP, seus funcionários e parceiros que com ele colaboram, o respeito e a dignidade a que têm direito.

4. Em suma

Na primeira parte deste trabalho, procurei relacionar o meu interesse pela linguística, pela política linguística e por uma hipotética diplomacia linguística, expondo alguns aspetos da minha história pessoal. Em seguida, fui dando conta das reflexões que me suscitaram os contactos profissionais que estabeleci com diplomatas da CPLP entre 2010 e 2022, fruto da minha colaboração com o IILP, o que justifica a atenção que, no texto, concedo ao Instituto a partir de então. Por fim, procurei entender o que poderia significar 'diplomacia linguística' e que áreas de atuação poderia abarcar. Por fim, procurei fornecer algumas sugestões de ação e atuação que possam contribuir para uma agenda para o diálogo em língua portuguesa.

Confesso a minha atração pelo projeto CPLP e, ainda mais, a minha confiança em que o IILP pode constituir um motor determinante do desenvolvimento da Comunidade e dos países que a constituem. Não posso, porém, deixar também de admitir o desencanto e a frustração que a ineficácia da CPLP e do IILP me provocam.

Uma primeira medida, simples e indolor, a adotar seria resolver a questão estatutária do IILP, levando a que os Estados-membros ratifiquem com urgência os Estatutos do Instituto

³² Considero que a adoção desta medida seria de elementar justiça e contribuiria para a estabilidade e o desenvolvimento quer do IILP, quer da própria CPLP.

Internacional da Língua Portuguesa, aprovados em 2010 e apenas ratificados por Portugal, para que o IILP possa funcionar com a estabilidade e a tranquilidade necessárias. Esta versão dos Estatutos, além de mais adequada à realidade que o hoje a língua portuguesa conhece, tem a virtude de traçar um elenco de atribuições gerais ao IILP, que, por si só, constituem uma agenda e um plano estratégico para o desenvolvimento da língua portuguesa e que por isso transcrevo:

Artigo 4.º

(...)

2 — São atribuições gerais do IILP:

- i) Elaborar e executar planos estratégicos e consequentes programas e projetos conducentes à internacionalização efetiva da Língua Portuguesa, nomeadamente junto de organizações internacionais e regionais;
- j) Promover, propor e apoiar a execução de projetos culturais, numa perspetiva integrada e de interculturalidade, de difusão da Língua Portuguesa;
- k) c) Fomentar a promoção do conhecimento da Língua Portuguesa e das Culturas da CPLP, nomeadamente, através do estabelecimento de redes de investigação e de intercâmbio de experiências;
- l) Promover e acompanhar o desenvolvimento de instrumentos de normalização linguística comum;
- m) e) Propor e ou avaliar projetos multilaterais de apoio ao desenvolvimento e otimização das competências em Língua Portuguesa em Estados membros da CPLP;
- n) f) Apoiar a adequada articulação entre a Língua Portuguesa e as demais línguas nacionais; g) Assegurar a representação da CPLP em fora internacionais sobre multilinguismo e multiculturalismo, por delegação do Secretário Executivo da CPLP;
- o) h) Colaborar com instituições dos Estados membros e de Estados terceiros no desenvolvimento de programas e projetos relevantes para a Língua Portuguesa.

Acredito que para muitos leitores, muito do que aqui postulo não passe de prefiguração de uma utopia. Os sucessos conseguidos na última década no IILP, apesar de todas as dificuldades, contrariedades e desentendimentos, se mais não houvesse, seriam para mim razões suficientes para acreditar em utopias.

“A utopia está lá no horizonte. Me aproximo dois passos, ela se afasta dois passos. Caminho dez passos e o horizonte corre dez passos. Por mais que eu caminhe, jamais alcançarei. Para que serve a utopia? Serve para isso: para que eu não deixe de caminhar.”

– Fernando Birri, citado por Eduardo Galeano em *Las palabras andantes*. Madrid: Siglo XXI, 1994³³.

³³ Fonte da citação: Revista Prosa Verso e Arte. <https://www.revistaprosaversoearte.com/para-que-serve-a-utopia-eduardo-galeano/>

5. Referências bibliográficas

Amparo, G. A. S.; Moreira, J. B. 2021. *A diplomacia não tem rosto de mulher: o Itamaraty e a desigualdade de gênero*. Meridiano 47, 22:e22001, 2021.

Gois, P.; Marques, J. C. 2018. *Retrato de um Portugal migrante: a evolução da emigração, da imigração e do seu estudo nos últimos 40 anos*. E-cadernos CES, 29.

Judd, E. L.; Tan, L.; Walberg, H. J. 2001. *Teaching Additional Languages*. s/l: UNESCO International Bureau of Education, International Academy of Education. Disponível em <https://unesdoc.unesco.org/ark:/48223/pf0000125455>.

Moita, L.; Pinto, L. V.; Pereira, P. (coords.). 2019. *Estudo da estrutura diplomática portuguesa*. Lisboa: OBSERVARE – Observatório de Relações Exteriores, Universidade Autónoma de Lisboa. DOI: <https://doi.org/10.26619/978-989-8191-97-7>

Serrão, J. 1970. *Conspecto da emigração portuguesa*. *Análise Social*, Vol. 8, No. 32, 597-617.

Španjević, M. 2014. *Diplomacia cultural: formas e possibilidades no terceiro setor*. Dissertação de Mestrado em Cultura e Comunicação. Lisboa: Faculdade de Letras da Universidade de Lisboa.

6. Anexo

Contextos de ocorrência da expressão “política linguística”, sublinhada.
Expressão de pesquisa “diplomacia linguística”+.pt

A. Revista Negócios Estrangeiros, publicação do Instituto Diplomático do Ministério dos Negócios Estrangeiros (Portugal), n. 21 / outubro de 2021, p. 217-218: «[Sérgio Vieira de Mello] Dividia as suas intervenções em duas ou três partes (salteando inglês, francês, português e tétum) num estranho jogo de diplomacia linguística que também era necessário fazer e que ele dominava melhor do que ninguém, nunca agradando inteiramente a nenhum dos lados.»

B. “Transcrição da história concedida pelo diplomata Luís de Oliveira Príncipe”, em HISTÓRIA DO FUTURO: Diagnóstico e perspectivas de políticas públicas para o ensino/aprendizagem de PLE-PL2 no Brasil do século XXI, dissertação de mestrado, UnB, Brasília, 2013, p. 184: “Eu respeito muito Portugal, falar mal de Portugal é, de certa forma, falar mal de nós mesmos e do ponto de vista linguístico temos o mesmo desafio, que é viver rodeados de países que falam o espanhol e que juntos são maiores do que nós. Portugal tem a Espanha e nós temos todos os países de língua espanhola ao redor, mais os da América Central e o México. Para estar à altura deste desafio, temos que ter uma diplomacia linguística, coordenada muito com Angola e muito com Moçambique.”

C. Língua, comunidade linguística e relações internacionais: uma análise pós-colonial da lusofonia, trabalho de conclusão de curso, UFRGS, 2017, p. 41: «Carneiro (2013, PLATÔ N.9 V.5 2022

p. 196) aponta a conexão entre diplomacia linguística e cooperação técnica e econômica na política externa brasileira do período: "A atuação internacional do Brasil através de convênios internacionais no âmbito acadêmico geridos pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) com países como o [...] Timor-Leste, a criação da [...] UNILAB, bem como a assinatura de diversos acordos de cooperação a serem geridos pela Agência Brasileira de Cooperação (ABC) nos países da CPLP revelam a articulação estreita de projetos que visam à promoção da língua portuguesa, com projetos que desenvolvem relações econômicas entre diferentes países ou ainda que visam a projeção do Brasil no plano internacional".»

D. "De todas as línguas se pode ver o mar: o Português e as línguas globais", artigo publicado na obra *O português na Casa do Mundo*, Hoje, Ed. Húmus, Braga, Portugal, p. 137: «Em línguas de poucos milhares de falantes, limitadas a uma região geográfica, eles são mais ou menos inofensivos, porque não envolvem questões de política linguística internacional. Mas em línguas de muitos falantes, globalmente dispersas, faladas em estados politicamente independentes podem acarretar incómodos e problemas de diplomacia linguística.»

E. Portugal, o país com mais "doutores" (e "engenheiros"), blogue *A Filosofia no Ensino Secundário* (<https://rolandoa.blogs.sapo.pt>), 2007: «Maria Velho da Costa lembra que, nos seus tempos de leitora de português no King's College em Londres, na década de 80, "os alunos tinham imensa dificuldade em perceber as várias fórmulas de address" existentes na língua portuguesa. Você, tu, o senhor, vossa excelência, senhor doutor, senhor professor, senhor engenheiro, excelentíssimo senhor, sôtor, senhora dona... Como explicar a um não-português a diplomacia linguística portuguesa, as suas possibilidades caprichosas? Como explicar-lhe, por exemplo, o sentido de chamar "engenheiro" a alguém que não pratica engenharia?»

F. Publicação da ONG AIDGLOBAL 2012: «Na opinião da Plataforma [Plataforma Portuguesa, que representa a maioria das Organizações Não Governamentais de Desenvolvimento (ONGD) portuguesas], "a Cooperação não pode ser subordinada à diplomacia linguística e cultural", criando um único Instituto da Cooperação e da Língua. "Só por si, a língua portuguesa não é fator de luta contra a pobreza, nem contribui para o desenvolvimento. O facto de os PALOP e Timor-Leste se encontrarem, quase na totalidade, no fim da lista dos países com mais baixo índice de desenvolvimento humano, só o comprova", afirma.»

G. Revista *Visão*, Portugal, 2018: «Donald Trump utilizou, mais uma vez, a sua vertente de performer que carrega nos adjetivos e despreza qualquer tipo de diplomacia linguística.»

H. Programa de leitorado: diálogo entre política linguística externa e formação de professores de PFOL no Brasil, tese de doutoramento, USP, 2017, p. 123 «Falar sobre o leitor significa discutir o perfil que ocupa lugar de centralidade no quadro das políticas ligadas à promoção e internacionalização da língua e da cultura do Brasil. Suas ações desempenham um papel de extrema importância na diplomacia linguística e cultural brasileira, seja no interior de quadros institucionais universitários e/ou em organismos oficiais como embaixadas.»

I. DIA NACIONAL COMEMORADO NO PORTO - França ascende à primeira posição do investimento estrangeiro em Portugal, artigo publicado na revista Vida Económica, Portugal, 2013, p. 25: «Sobre o Consulado-Geral de França no Porto, Aude de Amorim salientou um crescimento de 6% na comunidade inscrita, ultrapassando 5500 pessoas. Mas, de acordo com o Censos 2011, há mais de 20 mil franceses e franco-portugueses que residem na região Norte, ou seja, quatro vezes mais. Como prioridades para a ação na região Norte, Aude de Amorim referiu a diplomacia económica para apoiar e promover as empresas francesas que se instalam em número crescente em Portugal e reforçar a atratividade de França. Destacou também a diplomacia migratória, enquanto novo conceito que valoriza as vertentes da mobilidade, bem como a diplomacia linguística e cultural em torno dos valores da diversidade cultural e do multilinguismo, reforçando o interesse e o espaço da língua francesa.

J. Título de obra prefaciada por Jean-François Sirinelli, s/d, https://pt.frwiki.wiki/wiki/Jean-François_Sirinelli: «Política cultural francesa e diplomacia linguística: Alliance Française, 1883-1940, François Chaubet, Paris, L'Harmattan, 2006.»

K. Documento 7. "Quais as principais questões que se colocam à revisão da estratégia de cooperação e quais são os temas incontornáveis que deverão estar explanados no documento", artigo publicado no Relatório Aid Watch 2012 – Uma leitura da Cooperação Portuguesa desde 2003, 2012, p. 40: «Actualmente, persistem a ausência de rumo estratégico e a falta de transparência da política de Cooperação. As finalidades da Cooperação Portuguesa estão por clarificar, mas pode-se aferir quais são a partir de algumas medidas tomadas. A extinção do IPAD e a sua fusão com o Instituto Camões evidencia a opção de colocar a Cooperação ao serviço da diplomacia linguística e cultural, contrariando as recomendações do CAD/OCDE.»

Procedimentos teórico-metodológicos para recolha e sistematização de terminologias no âmbito da Comunidade de Países de Língua Portuguesa (CPLP)

Gladis Maria de Barcellos Almeida e Tanara Zingano Kuhn

Sobre o autor

Gladis Maria de Barcellos Almeida (Universidade Federal de São Carlos – UFSCar, docente e pesquisadora, doutora, doutoramento em Linguística e Língua Portuguesa, gladis@ufscar.br)

Tanara Zingano Kuhn (Centro de Estudos de Linguística Geral e Aplicada da Universidade de Coimbra – CELGA-ILTEC, pesquisadora, doutora, doutoramento em Linguística Aplicada, tanarazingano@outlook.com)

Resumo

As *terminologias* são os conjuntos das palavras e expressões específicas empregadas num dado domínio científico e técnico. Quando adequadamente sistematizadas, tornam-se um importante recurso político-linguístico, na medida em que favorecem a internacionalização da língua e seu emprego em organizações internacionais. Assim, atendendo às demandas do Tratado do Acordo Ortográfico da Língua Portuguesa de 1990, do Plano de Ação de Brasília (PAB) de 2010 e do Plano de Ação de Lisboa (PALis) de 2014, foi concebido o projeto *Terminologias Científicas e Técnicas Comuns* (TCTC), cujo objetivo geral foi a criação e disponibilização pública, na plataforma do Vocabulário Ortográfico Comum (VOC), de terminologias científicas e técnicas para os países da CPLP, com estrutura comum, geridas conjuntamente, gratuitas e de livre acesso. Para o desenvolvimento desse projeto, estabeleceu-se a base teórica, qual seja: uma Terminologia de orientação descritiva fundamentada em princípios da Linguística. Essa abordagem foi pertinente, já que as diferenças culturais entre os Estados-membros, que tanto enriquecem a língua portuguesa, deviam ser respeitadas e preservadas. Estabeleceram-se, também, as etapas metodológicas que guiaram as Equipes Nacionais (ENs), tanto para o alinhamento de termos entre as diferentes variedades do português, quanto para a elaboração de uma terminologia de raiz (do zero). Como um dos grandes resultados desse projeto, foi criada a plataforma TECER para facilitar o trabalho terminológico de todos os atores envolvidos no processo bem como para integrar os dados ao VOC.

Palavras-chaves

Língua portuguesa, terminologia, CPLP, TCTC, VOC.

1. Introdução: origem e relevância do projeto Terminologias Científicas e Técnicas Comuns (TCTC)

As *terminologias* são os conjuntos das palavras e expressões específicas (isto é, os termos) empregadas num dado domínio científico e técnico. Em usos especializados, a precisão terminológica é fundamental; os termos são a base da comunicação especializada. É, por isso, através da língua que se dão os intercâmbios tecnológicos e econômicos. Para os tradutores e intérpretes proporem equivalências noutras línguas, é preciso que existam, antes, terminologias adequadamente construídas.

As terminologias comuns são um recurso essencial para a internacionalização da língua portuguesa, para a interpretação e para o seu fortalecimento como língua de negócios, tendo um efeito multiplicador para a construção de outros recursos, nomeadamente computacionais, e para apoio à tradução. Trata-se de um instrumento essencial para emprego nas organizações internacionais por parte dos agentes de todos os países de língua oficial portuguesa.

É nas terminologias que existem algumas das maiores diferenças entre os países da Comunidade de Países de Língua Portuguesa (CPLP); ao mesmo tempo, é na terminologia que mais facilmente é possível coordenar políticas linguísticas entre os países, especialmente para a definição conjunta de termos para novos conceitos e realidades no domínio da ciência, tecnologia e indústria. Para isso, no entanto, é necessário criar instrumentos comuns que reúnam o material já existente, alinhem esse material, tornando-o compatível, e que se publique o resultado numa plataforma comum, tornando-o acessível ao público especializado e à indústria.

As Terminologias Científicas e Técnicas Comuns (TCTC) são um instrumento imprescindível para reforçar a realização de negócios em língua portuguesa e para a sua internacionalização, constituindo-se como uma ferramenta para a atividade de intérpretes e tradutores em organizações internacionais, ao permitirem identificar desarmonias existentes e pensar políticas conjuntas para as sanar nos instrumentos comuns criados no âmbito do Instituto Internacional da Língua Portuguesa (IILP), instituto que tem uma vocação explícita para articulação de esforços técnicos, científicos e financeiros dos Estados-membros (EMs), tanto para a promoção interna como para a promoção externa da língua portuguesa, desenvolvendo a sua atividade como um órgão colegiado dos países, numa perspectiva inovadora de gestão supranacional da língua.

O IILP, sob os auspícios da CPLP e para tal mandatado pelo *Plano de Ação de Brasília – PAB* (CPLP, 2010), liderou a elaboração do Vocabulário Ortográfico Comum da Língua Portuguesa (VOC), cuja plataforma foi oficialmente lançada em 2015 na sede da CPLP. Com o VOC já em pleno funcionamento, coube, também ao IILP, a elaboração, a partir do trabalho de base já levado a cabo no VOC, das *Terminologias Científicas e Técnicas Comuns*, conforme exigido no Tratado do Acordo Ortográfico da Língua Portuguesa e indicado no PAB (III, 4).

A existência de terminologias científicas e técnicas harmonizadas em todo o espaço da CPLP constitui:

- i) uma vantagem em termos económicos;
- ii) um importante instrumento de consolidação e reforço da língua portuguesa nas organizações internacionais a que pertencem países de língua oficial portuguesa;
- iii) um meio privilegiado para veicular informação e construir conhecimento na comunicação especializada. As instituições e os organismos responsáveis pela normalização terminológica nos países da CPLP deverão ser congregados em um projeto, para a elaboração, edição e distribuição de glossários, com áreas ou domínios cuja eleição deverá obedecer a critérios que atendam às realidades terminológicas dos EM. (CPLP, 2010, III, 4)

O mais recente *Plano de Ação de Lisboa – PALis* (CPLP, 2014), dando como prioritária a transferência de conhecimento entre EMs, reforça também “a importância da harmonização terminológica” (PALis, III, ii), e recomenda como medida:

[...] o reconhecimento do papel das terminologias científicas e técnicas como instrumento facilitador da comunicação entre os cientistas e investigadores dos países da CPLP e, sobretudo, como facilitador do ensino da ciência, da partilha do conhecimento e da disseminação da cultura científica, tendo em conta os projetos de normalização e/ou harmonização linguística já em curso (CPLP, 2014, III, i, 3).

Assim, o projeto Terminologias Científicas e Técnicas Comuns (TCTC)¹ teve como objetivo global a criação e disponibilização pública, na plataforma do VOC, de terminologias científicas e técnicas para os países da CPLP, com estrutura comum, geridas conjuntamente, no âmbito do IILP, gratuitas e de livre acesso, dando condições para que a língua portuguesa dispusesse de bases de dados terminológicos capazes de a colocar a par das línguas com mais recursos para a tradução e a interpretação, para emprego em organizações internacionais e para a criação de ferramentas derivadas, em particular no âmbito do processamento computacional da língua. Como resultado, previa-se a criação de recursos terminológicos conjuntos para todos os países de língua oficial portuguesa que fazem parte da CPLP.

Embora haja diversos recursos terminológicos para os países de língua portuguesa, na maioria das vezes, esses recursos:

- são executados setorialmente, para apenas um ou um conjunto reduzido de domínios científicos e tecnológicos;
- não são interligados nem feitos de acordo com a mesma metodologia, sendo por isso incompatíveis entre si;
- são pensados apenas para um dado país, sem a preocupação de articulação com os demais;
- em muitos casos estão desatualizados;
- em geral, não estão disponíveis ou apenas o estão em suporte-papel.

¹ As autoras agradecem enormemente a fundamental colaboração de José Pedro Ferreira e Margarita Correia no planeamento inicial do projeto Terminologias Científicas e Técnicas Comuns (TCTC), bem como na redação do projeto submetido ao Fundo Especial da CPLP em 2015.

As TCTC dão resposta a esses problemas integrando os recursos existentes numa plataforma única, o VOC, recurso do IILP já reconhecido pela CPLP, superando deste modo a necessidade e os custos inerentes de desenvolvimento de raiz de uma plataforma terminológica. Com a existência da plataforma comum, fica assegurada a publicação dos resultados e a sua difusão, dado que a plataforma existente cumpre já a função de distribuir dados de referência para a área da ortografia. A integração numa plataforma única elimina também o problema da incompatibilidade, pois os dados terão que ser conformados a um formato acordado e comum a todos os EMs. Por último, serão recuperados todos os dados existentes e utilizáveis, evitando-se ter que desenvolver de raiz todo o trabalho e evitando-se o desaproveitamento de recursos.

As TCTC devem incluir marcação de restrições geográficas e identificação do domínio científico e técnico a que diz respeito cada termo, assim como ligação explícita para a entrada de cada termo que esteja incluído no VOC, dando assim acesso aos usuários às suas propriedades formais (por exemplo: a ortografia, o plural ou a divisão silábica).

Nesse sentido, o projeto teve como objetivos específicos:

1. formar equipes nacionais (ENs) capazes de levar a cabo o trabalho de cada EM;
2. identificar, nos EMs, os repertórios terminológicos existentes;
3. criar meios tecnológicos para albergar esses repertórios;
4. obter e harmonizar os recursos obtidos, alinhando-os com as entradas do VOC, e ali disponibilizar os resultados ao público em geral e ao público especializado;
5. criar as fundações de uma rede multinacional de especialistas em terminologia nos EMs da CPLP que possa, no futuro, desenvolver terminologias de raiz e outras tarefas necessárias.

O projeto TCTC teve como públicos-alvo: os EMs da CPLP; os usuários da língua portuguesa; as instituições industriais, comerciais e científicas dos EMs; as organizações internacionais, especificamente os gestores de informação, os gestores de sistemas de informação e os tradutores e intérpretes a elas associados.

Como beneficiários finais, o projeto previa: o tecido produtivo, industrial e científico dos EMs da CPLP; os cidadãos dos EMs da CPLP e toda a comunidade internacional.

É importante ressaltar que todo o conhecimento técnico e/ou científico só pode ser ensinado, intercambiado, compartilhado e divulgado por meio de palavras. Essas palavras, adequadamente repertoriadas, vão constituir as terminologias de cada domínio das ciências e das técnicas. Todos os países da CPLP têm suas competências e habilidades (em diversos domínios) que os destacam no cenário internacional. Os países plantam, criam, confeccionam, fabricam bens, como também têm expertise para oferecer uma variedade enorme de serviços; em tudo isso há **terminologia**.

Entende-se que a função precípua do sistema lexical de uma língua natural é nomear o mundo, e o mundo inclui as ciências e as técnicas. Portanto, expressar o mundo sem terminologias não é expressar o mundo completo.

2. Princípios teóricos que nortearam o projeto TCTC

A palavra **terminologia** pode ter dois significados distintos. O primeiro refere-se ao conjunto de termos próprios de uma ciência, técnica, arte, segmento industrial ou atividade

profissional. Como exemplos, podem-se citar as terminologias da Nanotecnologia, da Instrumentação Cirúrgica, da Música, da Indústria Automotiva, da Enfermagem etc. O segundo significado refere-se não só ao conjunto de práticas e métodos utilizados na compilação, descrição, gestão e apresentação dos termos de uma determinada linguagem de especialidade (terminologia enquanto atividade) (SAGER, 1993), como também ao conjunto de postulados teóricos necessários para dar suporte à análise de fenômenos linguísticos concernentes à comunicação especializada, incluídos aí os termos, evidentemente (terminologia enquanto teoria) (ALMEIDA e CORREIA, 2008). Para efeito de clareza, emprega-se normalmente **terminologia**, com inicial minúscula, para designar vocabulário ou repertório vocabular; para terminologia enquanto atividade e teoria, emprega-se **Terminologia**, mas com a inicial maiúscula.

A prática de recopilar termos de uma dada área de conhecimento é muito mais antiga do que se pode imaginar, haja vista que o ato de nomear o mundo e as coisas sempre foi uma necessidade do homem. Segundo Barros (2004), tem-se registro dos primeiros repertórios terminológicos a partir de 2600 a.C., "feitos pelos sumérios em forma de tijolos de argila" (p. 29), nos quais havia palavras relacionadas "a profissões, gado, objetos comuns e divindades" (VAN HOOFF, 1998, *apud* BARROS, 2004, p. 29). A partir daí, várias iniciativas de construir repertórios terminológicos afluíram em distintos locais, mas é no século XVIII que vai ocorrer uma expressiva contribuição de recopilação de termos, notadamente nos domínios da Química, por intermédio de Lavoisier e Berthold, e da Botânica e Zoologia, graças ao incansável trabalho de Linné (CABRÉ, 1993).

O interesse dos especialistas de cada área pela Terminologia deveu-se à diligência desses cientistas para relacionar as denominações aos conceitos científicos. Isso se prolonga no século XIX, em que o desenvolvimento progressivo das ciências faz com que os cientistas busquem cada vez mais entender e descrever, sobretudo, as regras de formação dos termos de cada domínio de especialidade. Essa necessidade é expressa nos colóquios e/ou congressos internacionais de botânicos (1867), zoólogos (1889) e químicos (1892), ocorridos no final do século XIX (CABRÉ, 1993).

Na primeira metade do século XX, a necessidade já não é somente de relacionar denominações a conceitos, mas denominar conceitos novos e, principalmente, harmonizar as novas denominações. Isso porque se observa nessa época um progresso acelerado das ciências e das técnicas, como também um rápido desenvolvimento de tecnologias (CABRÉ, 1993). É nessa atmosfera que surgem os primeiros autores que se propuseram a sistematizar a prática terminológica: "E. Wüster (1898-1977), considerado o fundador da Terminologia moderna e principal representante da chamada Escola de Viena, e o russo D.S. Lotte (1889-1950), fundador da Escola Soviética de Terminologia" (CABRÉ, 1993, p. 22).

Embora para Rondeau (1983, *apud* CABRÉ, 1993) Lotte seja considerado "o verdadeiro pai da Terminologia enquanto disciplina científica" (CABRÉ, 1993, p. 22), Picht (1984, *apud* CABRÉ, 1993), vai defender que "a tese de doutoramento de Wüster é a primeira exposição de uma teoria da Terminologia" (CABRÉ, 1993, p. 22). A tese, publicada em 1931 e intitulada *A normalização internacional da terminologia técnica*, acaba por inaugurar a então chamada Teoria Geral da Terminologia (TGT), "concebida como disciplina autônoma e definida como um campo próprio de relação entre as ciências das coisas e outras disciplinas como a linguística, a lógica, a ontologia e a informática" (CABRÉ, 1996, p. 6). Delineia-se, então, o caráter multidisciplinar da Terminologia. A difusão da versão russa da tese de Wüster suscita maior interesse pela Terminologia nos domínios especializados e influencia a criação do

Comitê Técnico 37 Terminologia (TC37) da ISA (*International Standardization Association*) da Federação Internacional das Associações Nacionais de Normalizadores, a precursora da atual ISO (*International Standardization Organization*) (CABRÉ, 1993; CABRÉ, 1996).

A partir da segunda metade do século XX, a Terminologia passa a ser de interesse também dos linguistas, isso porque a disciplina deixou de ser vista apenas como um instrumento de normalização de termos, tornando-se mais um instrumento de comunicação (CABRÉ, 1993).

A teoria iniciada por Wüster é considerada o desenvolvimento teórico mais sistemático e coerente já realizado sobre os termos (CABRÉ, 1996). De acordo com Felber (1996), a linguagem técnico-científica funcionava como a *língua de uso* em oposição à *língua literária*. Razão pela qual ela tinha de ser eficaz e não conter ambiguidades. Nasce então a ideia da normalização da terminologia e da organização consciente da língua. Decorre disso o método que acabou tornando-se imprescindível na tecnologia: a normalização (FELBER, 1996).

O propósito de difundir terminologias normalizadas que contribuíssem para uma comunicação inequívoca regou todo o trabalho de Wüster e acabou dando à TGT a configuração de uma terminologia *representativa*, já que era preponderante denominar e etiquetar a informação, e *prescritiva*, pois as terminologias precisavam ser controladas para que a comunicação fosse inequívoca e eficaz, segundo Wüster.

De acordo com Cabré (1999), a TGT é uma teoria sistemática e coerente, válida para resolver um tipo de comunicação: a comunicação estandardizada; entretanto, essa teoria utiliza uma série de princípios que se mostram pouco satisfatórios no âmbito da comunicação real, já que a comunicação estandardizada é apenas uma das possibilidades da comunicação real.

É nesse sentido que, a partir dos anos 1980, começam a surgir críticas à TGT, pois, de acordo com alguns especialistas em Terminologia, a teoria de Wüster não permite descrever satisfatoriamente o léxico especializado. A pluralidade tipológica dos trabalhos motivada pelas distintas necessidades terminológicas, a dinâmica constante dos domínios especializados e, fundamentalmente, a diversidade das terminologias determinada pelas características pragmáticas da comunicação contribuem para que a TGT seja tomada como uma teoria insuficiente (CABRÉ, 1999).

Surgem, assim, perspectivas teóricas novas, sempre na tentativa de fazer da Terminologia uma disciplina mais descritiva (e menos prescritiva), utilizando os aportes teóricos da Linguística. Conforme explicita Kamikawachi (2009, pp. 17-18), seriam elas:

- I. *Socioterminologia* (GAMBIER, 1987, 1991; GAUDIN, 1993; BOULANGER, 1995): tem como foco a descrição dos termos em seus mais diferentes contextos de usos. A variação terminológica é estudada em movimentos sincrônico e diacrônico.
- II. *Terminologia de Base Textual* (HOFFMANN, 1998; CIAPUSCIO, 2003): tem como objetivo a compreensão minuciosa dos textos especializados, tanto em nível macro como microestrutural.
- III. *Teoria Comunicativa da Terminologia (TCT)* (CABRÉ, 1999, 2003): tem como princípio a visão linguística sobre a linguagem especializada. Nessa abordagem, o termo integra um determinado âmbito específico, sem perder as características próprias de qualquer unidade pertencente ao sistema linguístico das línguas naturais. Desta forma, um termo é considerado um signo linguístico em funcionamento numa situação de comunicação especializada.

IV. *Teoria Sociocognitiva da Terminologia* (TEMMERMANN, 2000): apresenta como destaque o papel dos modelos cognitivos, procurando mostrar as relações entre os processos de categorização e a linguagem. A abordagem é fundamentada na semântica cognitiva e questiona a centralidade da standardização em detrimento de uma descrição autêntica dos significados dos termos tal como aparecem nos textos.

No âmbito do projeto TCTC, procurou-se sistematizar um método de trabalho que, além de ser eficiente para a obtenção dos resultados esperados, estivesse em consonância com os postulados de uma Terminologia de orientação descritiva, fundamentada em princípios da Linguística. Filiar-se teoricamente a uma teoria descritiva significa, fundamentalmente, abandonar “o tratamento prescritivo das terminologias em favor de enfoques descritivos capazes de entender o léxico especializado como um elemento natural das línguas naturais” (KRIEGER e BEVILACQUA, 2005).

É esse, portanto, o grande pilar deste projeto, na medida em que as diferenças culturais entre os EMs, que tanto enriquecem a língua portuguesa, devem ser preservadas. Nas palavras de Gambier (1987, p. 319) “não há terminologia fora das práticas sociais, que são os discursos, fora de condições de enunciação definidas.”

3. Princípios metodológicos

No contexto do projeto TCTC, consideraram-se duas maneiras de obter repertórios terminológicos: i) alinhando terminologias já existentes nos países, tarefa que se denominou *harmonização*; ii) criando uma terminologia do zero, aqui denominada *terminologia de raiz*. Nas seções subsequentes, ambos os métodos serão apresentados e detalhados.

3.1. Harmonizando terminologias já existentes

Entende-se por harmonização o processo de alinhar termos diferentes, da mesma variedade ou não, mas que remetem ao mesmo conceito, de forma que todas as variações sejam consideradas, não havendo uma forma privilegiada em detrimento de outras, a não ser que uma EN assim o decida em relação às suas variações internas.

Do ponto de vista metodológico, a atividade de harmonização pressupõe que haja um repertório terminológico pronto e validado, ao menos por um país. Este repertório, então, pode ser organizado de duas formas²:

a) Quando se tem o equivalente em inglês de cada termo, a terminologia em inglês aparece na primeira coluna de uma grande planilha, e nas colunas subsequentes estão as variedades dos nove países da CPLP, como se pode observar na Figura 1;

b) Quando não se tem o equivalente em inglês, coloca-se na primeira coluna a definição do termo em português, de maneira a explicitar o conceito; na segunda coluna, a terminologia já validada de determinado país; e nas colunas subsequentes, os demais países, como apresentado na Figura 2.

² É preciso registrar que esse modo de trabalhar será facilitado a partir do momento em que a plataforma TE-CER (cf. Seção 4) estiver operante.

FIGURA 1
Planilha de alinhamento dos termos do Petróleo

termo em inglês	MODOS DE REFERIR EM LÍNGUA PORTUGUESA								
	variedade ANGOLA	variedade BRASIL	variedade CABO VERDE	variedade GUINÉ-BISSAU	variedade GUINÉ EQUATORIAL	variedade MOÇAMBIQUE	variedade PORTUGAL	variedade SÃO TOMÉ E PRÍNCIPE	variedade TIMOR LESTE
abandonment cap, corrosion cap	tampa de abandono	capa de abandono					tampa de abandono		
absorption plant		planta de absorção					instalação de absorção		
accelerator, set accelerator		acelerador de pega					endurecimento do cimento, endurecedor de		
acetylene welding		solda a acetileno					soldadura a acetileno		
acid bottle		bulbo de ácido					garrafa de ácido		
acoustic basement		embasamento acústico					soco acústico		
acoustic density log		perfil de densidade acústica					diagrafia de densidade acústica, log de densidade		
acoustic log	log acústico, log do som	perfil acústico					diagrafia acústica, diagrafia do som, log acústico, log do		
actinolite		actinolita					actinolite		
adapter		adaptador					adaptador do carretel de ligação		
adapter spool		carretel adaptador					carretel		

Fonte: elaborado pelas autoras

FIGURA 2
Planilha de alinhamento dos termos da Informática

Definição	variedade (já validada) BRASIL	variedade ANGOLA	variedade CABO VERDE	variedade GUINÉ-BISSAU	variedade GUINÉ EQUATORIAL	variedade MOÇAMBIQUE	variedade PORTUGAL	variedade SÃO TOMÉ E PRÍNCIPE	variedade TIMOR LESTE
interromper (um processo ou tarefa) voluntária ou acidentalmente a transferência de arquivos	abortar								
possibilidade de comunicar-se com um dispositivo, meio de armazenamento, unidade de rede, memória, registro, arquivo etc., visando receber ou fornecer dados	acesso								
reservar (um trecho de memória primária) para um programa, rotina ou armazenamento de dados	alocar								
profissional especializado em análise de sistemas	analista de sistema								
programa de proteção do computador que detecta e elimina os vírus (certos programas danosos) nele existentes, assim como impede sua instalação e propagação	antivírus								
programa de computador concebido para processar dados eletronicamente, facilitando e reduzindo o tempo de execução de uma tarefa do usuário	aplicativo								
ato de armazenar dados; armazenagem	armazenamento								
conjunto de dados digitalizados que pode ser gravado em um dispositivo de armazenamento e tratado como ente único	arquivo								
via de entrada em um sistema; ponto de entrada	atalho								
[ing.] ver cópia de segurança (forma aportuguesada: becape)	back-up								
Receber, através da rede de computadores, cópia de um arquivo localizado em máquina remota	baixar								
coleção ordenada de código de programas e rotinas, a que um programador pode recorrer para desenvolver outros programas	biblioteca								
[ing.] algarismo do sistema binário que somente pode assumir as formas 0 ou 1	bit								
que não admite saída ou entrada de dados, registros etc., ou que não os processa, sob determinadas circunstâncias (diz-se de unidade, programa ou terminal de computador); não disponível	bloqueado								
[ing.] o mesmo que <i>navegador</i>	browser								
[ing.] região de memória us. como área de armazenamento temporário de dados durante sua transferência entre dispositivos de diferentes taxas de transferência	buffer								
[ing.] defeito, falha ou erro no código de um programa que provoca seu mau funcionamento	bug								
[ing.] conjunto de bits adjacentes, ger. constituído por oito bits, que forma a unidade de informação, us. para representar um caractere	byte								
[ing.] área de memória onde é mantida uma cópia temporária de dados armazenados em um meio de acesso mais lento, com o objetivo de acelerar a recuperação dos dados	cache								

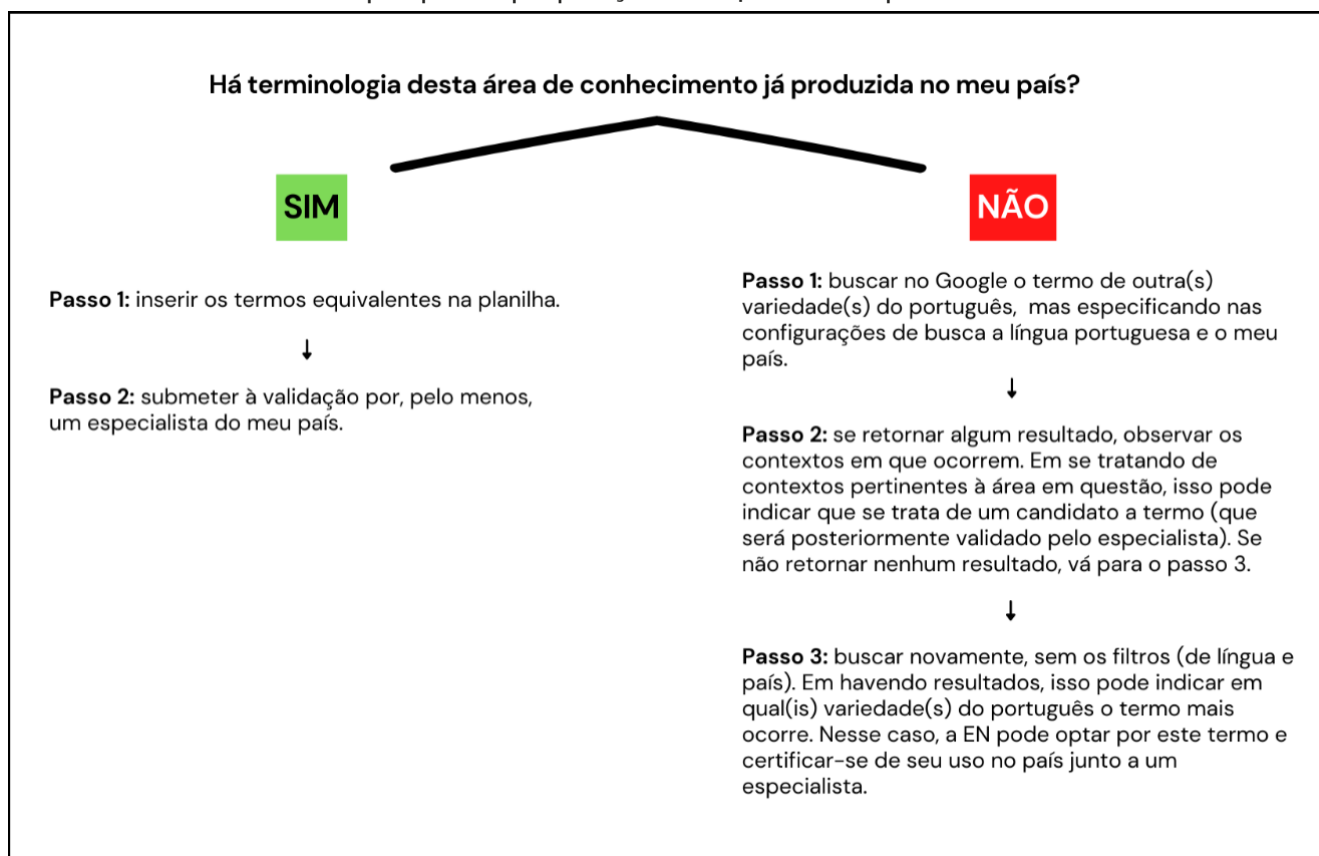
Fonte: elaborado pelas autoras

Tanto num caso como no outro, as planilhas foram encaminhadas a todos os países para que as respectivas ENs fizessem o alinhamento dos termos, inserindo nas colunas correspondentes os termos equivalentes em uso e validados nos seus países.

É importante ressaltar que a ação de alinhar um termo inclui: i) oferecer o(s) correspondente(s) terminológico(s) na variedade do país, quando houver; ii) aceitar formas já existentes em outras variedades, no caso de o país não possuir o(s) termo(s) correspondente(s).

Como os participantes das ENs geralmente não são especialistas nas áreas em que estão trabalhando, para a realização da harmonização (ou alinhamento) de termos entre variedades, sugeriu-se uma série de etapas para a proposição dos equivalentes, antes de submeter a lista de termos ao(s) especialista(s) de domínio para a validação. Essas etapas estão detalhadamente descritas no diagrama apresentado na Figura 3.

FIGURA 3
Etapas para a proposição de equivalentes pelas ENs



Fonte: elaborado pelas autoras

3.2. Fazendo terminologia de raiz

Organizar uma terminologia para uma área que não tenha ainda o seus termos identificados e repertoriados exige dedicação diligente, isso porque uma sequência de etapas metodológicas deve ser seguida e a ordem dessas etapas deve ser respeitada para que o resultado seja satisfatório.

Tendo em conta uma Terminologia de orientação descritiva (e não prescritiva), fundamentada em princípios da Linguística, as etapas são as seguintes:

1. delimitação da área de conhecimento;
2. seleção das fontes de onde provirão os textos que deverão compor o *corpus*³, bem como atenção aos requisitos recomendados e que já são consenso na Linguística de *Corpus*;
3. compilação dos textos, conversão de formatos, limpeza dos textos e nomeação dos arquivos;
4. anotação dos textos, se for o caso.

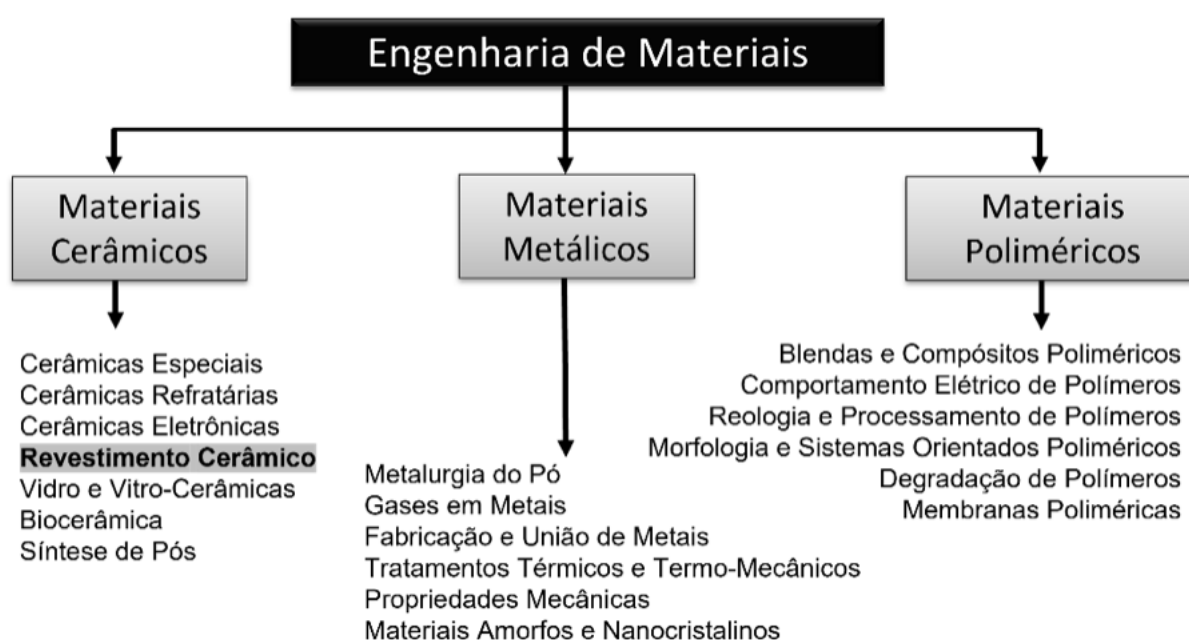
³ "Corpus é uma coleção de textos em formato eletrônico, selecionados de acordo com critérios externos para representar, tanto quanto possível, uma língua ou uma variedade linguística como fonte de dados para pesquisa linguística" (SINCLAIR, 2005).

A seguir, apresentar-se-á cada uma dessas etapas detalhadamente, chamando atenção para alguns aspectos que podem constituir-se em situações-problema e as suas possíveis soluções.

3.2.1. Delimitação da área de conhecimento

Quando se planeja realizar um projeto terminológico, deve-se delimitar o domínio, evitando-se eleger como objeto da pesquisa uma área inteira, como a Engenharia, a Medicina, a Psicologia etc., que constituem grandes áreas divididas em subáreas e, cada uma delas, por sua vez, desdobrando-se em outros níveis, cada um com as suas especificidades. Daí a necessidade de delimitar o domínio. Para ilustrar essa subdivisão em grande área, área e subáreas, observe-se a Figura 4 que retrata a grande área da Engenharia de Materiais, que já é um tipo de Engenharia.

FIGURA 4
Delimitando o domínio na área de Engenharia de Materiais



Fonte: Almeida (2000)

Conforme demonstra a Figura 4, a Engenharia de Materiais está dividida em três subáreas: Materiais Cerâmicos, Materiais Metálicos e Materiais Poliméricos. Cada subárea, por sua vez, divide-se em outras tantas especificidades, consistindo em distintas linhas de investigação. Cada especificidade destas tem suas próprias características no que concerne a matéria-prima, equipamento, procedimento, produto; transformando-se em vários universos conceituais, com muitos termos próprios. É mais proveitoso, portanto, eleger um dessas especificidades, como a de Revestimento Cerâmico, conforme assinalado na Figura 4, e torná-la o objeto da recolha terminológica.

Assim, lidar com uma grande área como um todo pode revelar-se contraproducente por várias razões:

- a) geralmente, as áreas se compõem de subáreas com distintas especificidades e geram um universo muito grande de fontes de obtenção dos textos que deverão compor o *corpus*. Além disso, há grande possibilidade de dispersão, que pode resultar num problema no momento da extração dos candidatos a termos⁴, pois com essa profusão de especificidades temáticas, corre-se o risco de deixar de considerar termos relevantes para determinada especialidade e fazer o inverso com outra, ou seja, acabar privilegiando uma em detrimento de outra;
- b) torna-se necessário contar com uma assessoria especializada muito maior, o que dificulta o trabalho. Além disso, a não delimitação do domínio pode acarretar um desequilíbrio entre as subáreas;
- c) frequentemente as subáreas são muito distintas, o que dispensa ter num mesmo produto temas tão diferentes.

Alguns fatores que podem auxiliar na delimitação do domínio:

- a) interesse dos especialistas da área em ter sua terminologia organizada e repertoriada;
- b) número de profissionais colaboradores com os quais se poderá contar;
- c) relevância de determinada especificidade do ponto de vista econômico, social, político, cultural, científico e/ou tecnológico para o país;
- d) facilidade de obtenção de textos já em formato digital para agilizar a compilação do *corpus*.

3.2.2. Seleção das fontes de onde provirão os textos

Escolhido o subdomínio com o qual se vai trabalhar, passa-se, então, a recolher informações para melhor conhecê-lo, para, a partir daí, chegar-se às fontes confiáveis de onde proverá o *corpus*.

Assim, é necessário realizar pesquisas (que podem ser feitas pela internet, utilizando-se motores de busca como o Google) para conhecer quais são as instituições e/ou entidades públicas, privadas ou mistas que têm tradição no referido subdomínio; quais universidades, laboratórios e/ou empresas realizam pesquisas nos temas referentes à especialidade escolhida; se todos esses organismos possuem material escrito; se este material está disponível e em formato digital.

É recomendável que toda essa pesquisa sobre as fontes seja referendada por um (ou mais) especialista de domínio, com o objetivo de atestar a confiabilidade das fontes.

Confirmadas as fontes, deve-se definir o tipo de *corpus*. Para isso, são levados em consideração os seguintes aspectos⁵:

- Que produto terminológico será gerado: apenas uma lista de termos sem definição? Quantos termos? Um glossário com entradas, definições e exemplos de uso? Uma base

⁴ Os candidatos constituem itens léxicos que se comportam nos seus respectivos contextos como termos, mas cuja autenticidade será validada posteriormente.

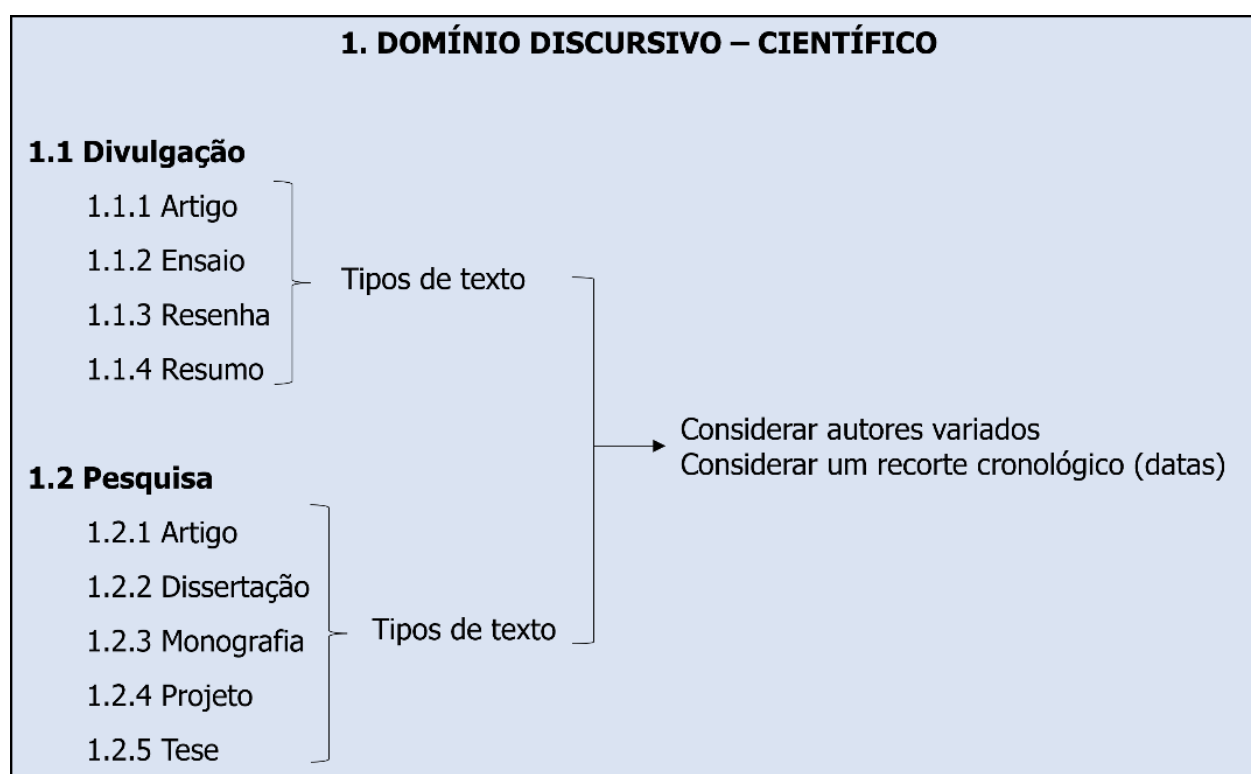
⁵ Evidentemente o objetivo principal aqui é elaborar terminologias que integrem o VOC, mas seria um grande desperdício de tempo e de recursos humanos não aproveitar todo esse trabalho empreendido por cada EN e não construir seus próprios produtos terminológicos (glossários, dicionários, ontologias etc.).

terminológica contendo um conjunto de fichas com campos predeterminados? Uma ontologia? Será um produto monolíngue, bilíngue, multilíngue ou monolíngue com equivalências?

- Qual será o público-alvo: especialistas de domínio? Público semiespecializado, como estudantes universitários, por exemplo? Público leigo? Tradutores, bibliotecários, cientistas da informação? Ou o projeto terminológico prevê todos esses perfis de consulente ao mesmo tempo?

Esses aspectos ajudarão a definir as seguintes características do *corpus*: tamanho, diversidade de domínios discursivos, diversidade de textos dentro desses domínios, diversidade de autores e recorte temporal. Para ilustrar, observe-se a Figura 5.

FIGURA 5
Tipos de texto do domínio discursivo - científico



Fonte: elaborado pelas autoras, com base em Souza (2011)

Segundo Sinclair (2005), o *corpus* deve ter o tamanho adequado ao tipo de pesquisa que se vai realizar e à metodologia a ser adotada na pesquisa. Quando se fala em tamanho de um *corpus*, não se trata somente do número total de ocorrências (*tokens*) e de palavras diferentes (*types*); trata-se de determinar com quantas categorias (domínios discursivos, tipos de textos, autores, datas etc.) um *corpus* deve contar, quantas amostras de cada categoria e quantas ocorrências existem dentro de cada amostra (PEARSON, 1998; KENNEDY, 1998).

Há uma diversidade de domínios que podem ser considerados, como científico (incluindo divulgação e de pesquisa), informativo, referencial, instrucional, informativo, técnico-administrativo e/ou oficial, jurídico etc. Numa pesquisa terminológica, sugere-se a

inclusão de, pelo menos, dois domínios (por exemplo: científico e instrucional). O Apêndice 1 traz uma relação bastante extensa desses domínios e seus tipos de texto, o que pode auxiliar na identificação das fontes para a elaboração do *corpus*.

Em relação à diversidade de textos, é importante considerar vários títulos distintos. Por exemplo, é possível que haja uma coleção completa on-line de uma importante revista científica e toda a sua compilação resultar num *corpus* de um milhão de palavras, o que é um tamanho razoável. Entretanto, o *corpus* perderá em diversidade, posto que contempla apenas uma revista.

No que se refere à diversidade de autores, é importante também diversificar, pois se o *corpus* se basear apenas na obra completa de dois ou três autores, corre-se o risco de muitas escolhas terminológicas serem apenas um traço estilístico dos autores selecionados.

Selecionados os domínios, os tipos de textos e os autores, é preciso considerar também um recorte cronológico, daí a importância das datas de publicação dos textos.

Além de todos esses aspectos mencionados, há que se levar em conta os requisitos recomendados e que já são consenso na Linguística de *Corpus*, tais como: autenticidade, representatividade, balanceamento, amostragem, diversidade e tamanho⁶ (cf. KENNEDY, 1998; PEARSON, 1998; BIBER, 1998 et al.; RENOUF, 1998; BERBER SARDINHA, 2004; SINCLAIR, 2005).

3.2.3. Compilação dos textos, conversão de formatos, limpeza dos textos e nomeação dos arquivos

Depois da seleção das fontes e dos textos, considerando todos os aspectos arrolados, inicia-se a compilação dos textos propriamente dita. A compilação consiste no armazenamento em arquivos predeterminados de todos os textos selecionados.

Podem-se buscar textos provenientes da *web* ou mesmo textos impressos. Para o caso de se utilizar a *web* como fonte para a compilação do *corpus*, as máquinas de busca são bastante úteis, como o *Google*, bastando utilizar palavras-chave escolhidas para a pesquisa em foco. Os textos impressos também podem ser utilizados, entretanto, será necessário digitalizá-los.

A conversão dos formatos *doc*, *docx*, *html*, *pdf* etc. para *txt* pode ser manual ou automática. Na conversão manual, utiliza-se o recurso copia-e-cola ou “salvar como” e optar por salvar como texto simples (*plain text*). Para a conversão automática, pode-se fazer uso de ferramentas próprias para isso. Algumas ferramentas frequentemente utilizadas para esse processo são o Apache Tika⁷ e pacotes em Python como o doc2text⁸ e o PyPDF2⁹. Os textos são salvos em seu formato puro, isto é, apenas as sequências de caracteres, sem marcações como negrito, itálico etc. Este formato normalmente adota a extensão *txt* e é conhecido como texto puro, texto simples ou *plain text*. NotePad++, Sublime Text são exemplos de aplicativos frequentemente utilizados para editar textos de *corpora*.

⁶ Informações mais detalhadas sobre a elaboração de um *corpus* e seus requisitos podem ser obtidas em Ferreira e Almeida (2012).

⁷ <https://tika.apache.org/>

⁸ <https://pypi.org/project/doc2text/>

⁹ <https://pypi.org/project/PyPDF2/>

A limpeza e formatação têm o objetivo de preparar o *corpus* para o processamento computacional, o que significa tirar imagens, gráficos, tabelas, números de páginas e demais anotações que não fazem parte do texto propriamente dito. A limpeza e a formatação possibilitam o processamento do *corpus* por ferramentas computacionais, tais como: contador de frequência, concordanciador, ferramenta de extração automática de termos etc.

A nomeação dos arquivos é também uma etapa importante porque confere organização ao conjunto de textos compilados. Sugere-se que a nomeação seja transparente, remetendo o pesquisador imediatamente para o meio de divulgação, o domínio discursivo, a fonte, a data de publicação, nessa ordem. Assim, é possível padronizar os nomes dos arquivos por meio de siglas que contenham caracteres alfanuméricos. Observe-se, a título de exemplo, a nomeação de um dos arquivos do projeto NanoTerm¹⁰: *IN-CD-INOVATEC-nanotec-05jul06*, significando:

IN: texto divulgado pela Internet

CD: gênero científico de divulgação

INOVATEC: fonte de onde o texto foi obtido

nanotec: caderno/seção em que o texto foi publicado na fonte

05jul06: data de publicação

Quando se tem um *corpus* com mais de 500 arquivos de texto, esse procedimento passa a ser imprescindível para o reconhecimento dos arquivos e a sua organização em pastas. Muitas vezes, as informações extratextuais que se deseja manter de um texto são muitas. Nesses casos, opta-se por preservar tais informações em um documento externo ao texto em si, ou utilizando marcações XML (cf. seção 3.2.4).

3.2.4. Anotação de textos

Em relação à anotação, são dois basicamente os níveis de representação das informações presentes num *corpus*: a anotação estrutural (também conhecida como marcação) e a anotação linguística.

A anotação estrutural do texto é um processo pelo qual partes de texto que constituem informações diferenciadas por sua relevância (ou pela falta dela)¹¹ são marcadas por etiquetas, geralmente seguindo a linguagem de marcação *Extensible Markup Language* (XML)¹². Esse tipo de anotação compreende a marcação de dados externos e internos dos textos. Como dados externos, entende-se a documentação do *corpus* na forma de um cabeçalho que inclui os metadados textuais, isto é, dados bibliográficos comuns, como tamanho do arquivo,

¹⁰ Projeto coordenado por Gladis Maria de Barcellos Almeida e intitulado Terminologia em língua portuguesa da nanociência e nanotecnologia: sistematização do repertório vocabular e elaboração de dicionário-piloto. Foi desenvolvido entre os anos de 2006 e 2008 no Grupo de Estudos e Pesquisas em Terminologia (GETerm), sediado no Departamento de Letras da Universidade Federal de São Carlos (UFSCar), São Carlos, SP, Brasil, com financiamento do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

¹¹ Por exemplo: título, resumo e legendas constituem informações relevantes; enquanto *abstract* e referências bibliográficas constituem informações irrelevantes, no caso de tratar-se de um projeto terminológico em língua portuguesa, por exemplo.

¹² Código de formatação próprio da Computação que tem como função inserir uma marcação específica para cada trecho do texto de forma que a especificidade desse trecho possa ser reconhecida por outros programas de Processamento de Linguagem Natural (PLN).

autoria, tipologia textual, data de publicação etc. Como dados internos, existe a anotação de segmentação do texto cru, que envolve: a) marcação da estrutura geral – capítulos, parágrafos, títulos e subtítulos, notas de rodapé e elementos gráficos como tabelas e figuras, e b) marcação da estrutura de subparágrafos – elementos que são de interesse linguístico, tais como sentenças, citações, palavras, abreviações, nomes, referências, datas e ênfases tipográficas do tipo negrito, itálico, sublinhado etc. Essas informações facilitam a posterior recuperação do texto bem como a geração de subcorpus, ou seja, podem-se selecionar todos os textos de determinado autor, ou de determinada época, ou de determinado gênero etc. (cf. ALUÍSIO e ALMEIDA, 2006).

A anotação linguística pode ser em qualquer nível que se queira, isto é, nos níveis morfosintático, sintático, semântico, discursivo, retórico etc., sendo inserida de três formas: manualmente, automaticamente (por ferramentas de PLN), ou semiautomaticamente (correção manual da saída de outras ferramentas) (ALUÍSIO e ALMEIDA, 2006). Há que se ressaltar, no entanto, que a anotação não é uma etapa obrigatória, é possível realizar uma pesquisa terminológica baseada em um *corpus* não anotado.

No que se refere à anotação linguística nos níveis morfológico e lexical, podem-se utilizar as seguintes opções gratuitas: TreeTagger, Freeling e Spacy.

3.3. Extração dos termos

A extração de candidatos a termos depende diretamente do *corpus* compilado, pois é nos textos que se observam os termos em uso, e é o uso que ratifica a relevância de determinado termo:

Os textos especializados fornecem o certificado dos termos, ou seja, uma prova de que eles existem e de que são efetivamente utilizados pelos especialistas. Além disso, eles informam a frequência de emprego de um termo (L'HOMME, 2004, p.120).

Há ainda outro importante aspecto a ser considerado durante a extração terminológica. Invariavelmente, os manuais de Terminologia postulam a designação termo para aquilo que, em língua geral, nomeamos palavra. Essa dicotomia acaba sugerindo que há diferenças substanciais entre essas unidades, o que poderia orientar métodos de extração semiautomática de termos (sobretudo aqueles baseados em conhecimento linguístico e dependentes de língua). Entretanto, não é o que se verifica quando se examinam diferentes repertórios terminológicos, já que: i) nem todas as terminologias possuem os mesmos padrões morfolexicais; ii) na maioria das vezes, os critérios que distinguem termo de palavra não residem nos aspectos formais da língua, mas sim nos aspectos semânticos e enunciativos; e, finalmente, iii) não existe um conjunto de termos isolados constituindo uma língua marginal à língua geral, o que há são signos linguísticos da língua natural que se realizam ora como palavras, ora como termos. Nesse sentido, o *corpus* tem papel extremamente relevante, porque em muitas terminologias, sobretudo naquelas cuja maioria dos termos é proveniente da língua geral¹³, apenas observando os signos linguísticos em contexto é que se pode, de fato, identificar se se trata de termo ou de palavra.

¹³ Observem-se, por exemplo, os termos agitador, barriga, bolha, brilho, coração negro, fileira, forno, mancha, mofo, orelha, peneira, risco, verruga etc. retirados da terminologia brasileira de Revestimento Cerâmico (ALMEIDA et al. 2011).

A extração de termos diz respeito à obtenção do repertório terminológico que se deseja construir. Há três abordagens por meio das quais se pode fazer a extração (ALMEIDA et al., 2006): i) a primeira, denominada estatística, utiliza-se de sistemas baseados em estatística; ii) outra abordagem existente é a linguística, em que os sistemas detectam padrões recorrentes de unidades terminológicas complexas, tais como “substantivo–adjetivo” e “substantivo–preposição–substantivo”, por exemplo; iii) o terceiro tipo é a híbrida, em que os sistemas começam a detectar algumas estruturas linguísticas básicas, tal como expressões nominais, e depois de os termos candidatos terem sido identificados, alguma estatística relevante é usada para decidir se eles correspondem a um termo. O inverso também é possível, começando-se com uma lista de candidatos levantados estatisticamente, sendo que a informação linguística, nesse caso, é usada para filtrar termos válidos da lista.

Na ausência de ferramentas linguísticas ou estatísticas, ainda assim é possível fazer a extração, pois vários programas¹⁴ de manipulação automática de textos vêm com estatísticas próprias, bem como oferecem buscas por expressões fixas. Além disso, as telas do concordanciador (ferramenta que apresenta uma lista exaustiva de todas as ocorrências e seu contexto), resultantes de uma busca realizada no *corpus*, sempre ajudam o pesquisador a identificar em contexto se determinado item léxico pode ser um termo. Por exemplo, ao buscar pelo padrão “substantivo + adjetivo” no Corpus de Português Escrito em Periódicos (CoPEP) (KUHN; FERREIRA, 2020) disponível na ferramenta de *corpus* CQPWeb¹⁵, obtemos uma lista (ou linhas de concordância) com todas as ocorrências no *corpus*, como se pode ver na Figura 6.

FIGURA 6
Tela da ferramenta de corpus CQPWeb

Your query "[pos = \"NOUN.*\"] [pos = \"ADJ.*\"]" returned 2,193,570 matches in 9,900 different texts (in 48,506,519 words [9,900 texts]; frequency: 45,222.17 instances per million words), ordered randomly [0.568 seconds]			
< << >> > Show Page: 1 Line View Show in corpus order Choose action... Go!			
No.	Text	Solution 1 to 50	Page 1 / 43872
1	8547	qualquer actividade desportiva durante a sua viagem , mas em que a	prática desportiva
2	3119	e lógica . No mosaico de percepções sobre o indigenismo , a	comunidade epistêmica
3	91	qualquer representação que faça essa reivindicação o faz com base em certas	noções específicas
4	9235	) . Não foram observadas quaisquer relações entre a TFG e o	aporte proteico
5	5836	uma vez que , a exemplo do que acontece em um	orçamento financeiro
6	7435	as ideias produzidas são os próprios afetos passivos ; e elas são	ideias inadequadas
7	6908	não ter realizado algum pedido de seus devotos , prática comum nas	crenças populares
8	3554	projectiva . O sujeito , em vez de pensar , coloca os	elementos assimbólicos
9	1584	3 , respetivamente . Para obtenção do índice final realizou -se a	soma simples
10	8388	e outras perdas , estes estão diretamente relacionados com as concentrações de	preparados enzimáticos
11	8841	são eles nossos aliados naturais em qualquer esforço tendente a mudar o	sistema normativo
12	4139	que as participantes em programas de prevenção demonstrem efeitos negativos do seu	desenvolvimento sexual
13	7499	respeito às decisões familiares quanto à educação , processo diretamente ligado às	características socioeconômicas
14	9171	Rh(III) para redução eletrocatalítica de prótons19,20 , em processos de hidrogenação ou	redução fotoquímica
15	4642	as seguintes possibilidades verdadeiras : Possibilidades factuais : Possibilidades deonticas : sangue	luvas não-sangue
16	6571	sejam dirigidos para elas ; se não , poderão ser dirigidos para	fontes distantes-atraves
17	5686	contribuirão para uma via fisiopatológica comum que em última análise resulta na	fibroproliferação tecidular
18	3018	Este caso parece portanto contrariar as tendências da evolução ideológica dos	partidos socialistas
19	7939	intima relação com o movimento preventivista de Gerald Caplan(2) , objetivando desenvolver	medidas concretas

Fonte: Kuhn e Ferreira (2020)

¹⁴ Um programa bastante amigável e gratuito é o Antconc (cf. em <https://www.laurenceanthony.net/software/antconc/>). Além disso, pode-se assistir aos tutoriais, cujos links se encontram nessa mesma página.

¹⁵ <https://ola.unito.it/CQPweb32/>

3.4. Elaboração da estrutura conceitual, inserção dos termos e sua validação

Os termos obtidos no processo de extração devem ser inseridos na estrutura conceitual, por isso ela deve ser organizada preliminarmente, ou concomitantemente à extração dos termos, já que, à medida que os termos vão sendo obtidos, pode-se ter uma visão real de quais serão os campos semânticos que deverão integrar a estrutura de conceitos (ALMEIDA *et al.*, 2006).

Para organizar a estrutura conceitual, primeiro se estabelece a hierarquização da área-objeto com as suas subáreas constitutivas e, dentro dessas subáreas, os campos semânticos pertinentes, no interior dos quais cada termo será alocado (cf. Figura 7).

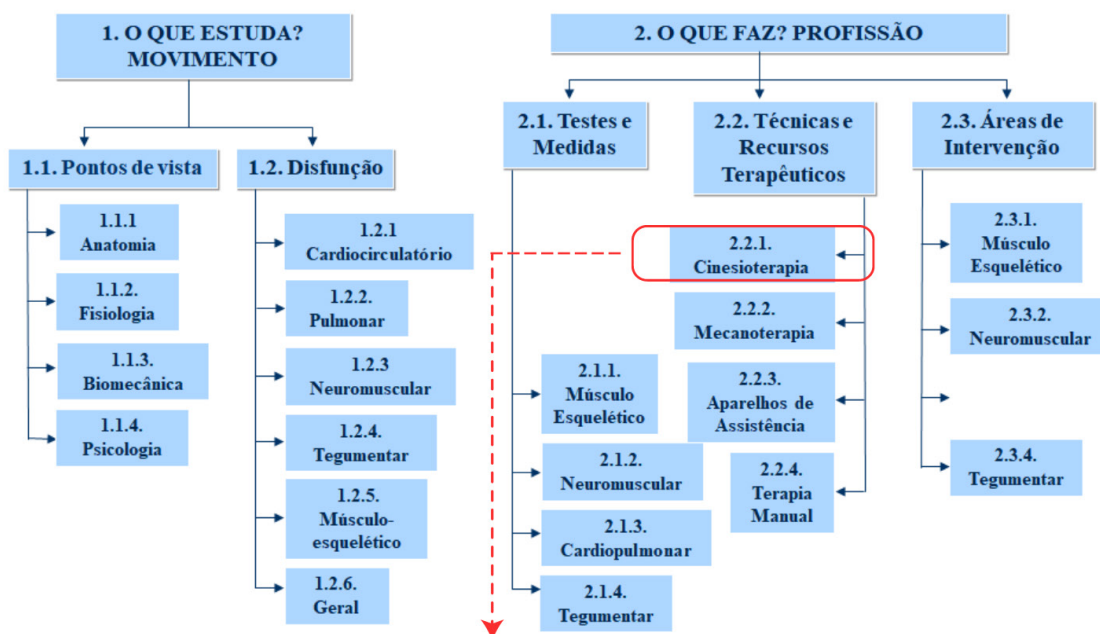
A estrutura deve ser elaborada pelos terminólogos com assessoria dos profissionais do domínio. Na pesquisa terminológica, a organização da estrutura conceitual é fundamental para: i) possibilitar uma abordagem mais sistemática de um campo de especialidade; ii) circunscrever a pesquisa, já que todas as ramificações da área-objeto, com seus campos, foram previamente consideradas; iii) delimitar o conjunto terminológico; iv) determinar a pertinência dos termos, pois separando cada grupo de termos pertencente a um determinado campo, poder-se-á apontar quais termos são relevantes para o trabalho e quais não são; v) prever os grupos de termos pertencentes à área-objeto, como também os que fazem parte de matérias conexas (ALMEIDA, 2000; ALMEIDA *et al.*, 2007). Para ilustrar, a Figura 7 traz a estrutura conceitual da Fisioterapia¹⁶, com os termos já inseridos no campo semântico Cinesioterapia. Essa estrutura conceitual com alguns termos inseridos foi organizada no Canva¹⁷, mas também pode ser elaborada no CmapTools¹⁸, ferramenta bastante amigável que permite elaborar estruturas conceituais de forma gráfica, permitindo a inserção de relações conceituais entre os termos, como se pode ver nos exemplos disponíveis no Anexo 1.

¹⁶ Projeto coordenado por Gladis Maria de Barcellos Almeida e intitulado Dicionário de Fisioterapia (DicFisio). Foi desenvolvido entre os anos de 2001 e 2004 no Grupo de Estudos e Pesquisas em Terminologia (GETerm), sediado no Departamento de Letras da Universidade Federal de São Carlos (UFSCar), São Carlos, SP, Brasil, com apoio financeiro (bolsas) do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

¹⁷ Canva é uma ferramenta gratuita de design gráfico on-line que pode ser utilizada para criar posts para redes sociais, apresentações, cartazes, vídeos, gráficos, diagramas, fluxogramas etc. Na versão paga, mais recursos de edição são oferecidos (https://www.canva.com/pt_br/).

¹⁸ Pode ser obtida gratuitamente aqui: <https://cmap.ihmc.us/> (acesso em: 14 nov 2021).

FIGURA 7
Estrutura conceitual da Fisioterapia



2.2. TÉCNICAS E RECURSOS TERAPÊUTICOS

2.2.1. Cinesioterapia

- 2.2.1.1. exercícios de alongamento
 - 2.2.1.1.1. alongamento passivo
 - 2.2.1.1.2. alongamento seletivo
 - 2.2.1.1.3. alongamento cíclico
 - 2.2.1.1.4. autoalongamento
- 2.2.1.2. Bobath
- 2.2.1.3. cadeia cinética
 - 2.2.1.3.1. cadeia cinética aberta
 - 2.2.1.3.2. cadeia cinética fechada
- 2.2.1.4. condicionamento físico
- 2.2.1.5. exercício terapêutico
- 2.2.1.6. exercício ativo
 - 2.2.1.6.1. exercício ativo-livre
 - 2.2.1.6.2. exercício ativo-assistido
 - 2.2.1.6.3. exercício ativo-resistido
- 2.2.1.7. exercício isocínético
- 2.2.1.8. exercício isométrico
- 2.2.1.9. exercício isotônico
- 2.2.1.10. exercício passivo
- 2.2.1.11. exercício excêntrico
- 2.2.1.12. exercício intermitente
- 2.2.1.13. exercício isométrico resistido
- 2.2.1.14. exercício de estabilização
- 2.2.1.15. exercício com resistência manual
- 2.2.1.16. decúbito
 - 2.2.1.16.1. decúbito supino
 - 2.2.1.16.2. decúbito prono
 - 2.2.1.16.3. decúbito lateral
- 2.2.1.17. diagonal funcional
- 2.2.1.18. diagonal primitiva
- 2.2.1.19. drenagem
 - 2.2.1.19.1. drenagem brônquica
 - 2.2.1.19.2. drenagem linfática
 - 2.2.1.19.3. drenagem pleural
 - 2.2.1.19.4. drenagem postural

A partir do momento em que os termos estão alocados na estrutura, pode-se proceder à sua validação pelos especialistas.

A validação dos candidatos a termos pode ser feita das seguintes formas:

1. pela comparação da lista de candidatos com uma lista de itens léxicos provenientes de um *corpus* de referência (*corpus* da língua geral);
2. pela submissão da lista de candidatos à análise de especialista(s) do domínio, para tanto, selecionam-se da estrutura conceitual determinados campos semânticos e pede-se que cada assessor assinale os termos considerados semanticamente relevantes em cada campo. A esse modo de selecionar termos denomina-se *critério semântico* (ALMEIDA *et al.*, 2007). É possível que, durante essa tarefa, o especialista observe a ausência de determinado termo. Nesse caso, ele mesmo pode fazer a inserção. Afinal, os termos foram obtidos no *corpus* e, por mais criteriosa que tenha sido a elaboração de um *corpus*, ele nunca é completo;
3. pela utilização dos dois procedimentos sequencialmente, ou seja, comparam-se as listas e, após a comparação, submete-se o resultado à apreciação do(s) especialista(s). Embora seja sempre recomendada a análise de especialista(s), muitas vezes não se opta por esta última possibilidade por questões de tempo, haja vista que esses profissionais normalmente não têm disponibilidade, atrasando, portanto, o decurso do projeto.

Depois de encerradas todas essas fases, a terminologia está pronta para ser enviada à Equipe Central (EC)¹⁹ para ser inserida na plataforma TECER (cf. seção 4). Recomenda-se que o repertório terminológico seja enviado como planilha do Excel.

4. Plataforma TECER

A plataforma TECER foi pensada e construída de modo a atender os objetivos do projeto TCTC (cf. Seção 1). Apesar da existência de ferramentas ou plataformas que permitem a criação e gerenciamento de terminologias (cf. ALMEIDA; OLIVEIRA, 2013 e SUHARDIJANTO; DINAKARAMANI, 2017), nenhuma atende a todas as necessidades em relação a esses objetivos.

Para facilitar a formação de equipes nacionais, é preciso uma ferramenta com uma interface amigável e em português, diminuindo, assim, a curva de aprendizado e agilizando o processo de formação. Um programa que não exija instalação e possa ser facilmente acessado via um navegador de internet também oferece um melhor ambiente para a formação das equipes.

A identificação de repertórios terminológicos existentes nos EMs não é suficiente para imediata inserção dos dados na plataforma. Esses dados precisam ser extraídos, limpos e formatados de modo adequado. Para tal, um conjunto de pacotes para extração de conteúdo da *web*, de *pdf*, entre outros formatos, está em constante desenvolvimento e aperfeiçoamento. Isto porque textos são submetidos nas mais variadas formas e não há como prever quais padrões serão encontrados.

¹⁹ A Equipe Central do projeto TCTC foi constituída entre os anos de 2015 e 2017 por: Gladis Barcellos Almeida (Brasil), José Pedro Ferreira (Portugal) e Margarita Correia (Portugal). Entre 2018 e 2020 por: Andressa Gomide (Brasil), Carla Maciel (Moçambique), Gladis Barcellos Almeida (Brasil) e Tanara Zingano Kuhn (Brasil). Entre 2021 e 2022 por: Abigail Tiny Cosme (São Tomé e Príncipe), Adelaide Monteiro (Cabo Verde), Andressa Gomide (Brasil), Carla Maciel (Moçambique), Gladis Barcellos Almeida (Brasil) e Tanara Zingano Kuhn (Brasil).

Para abrigar os repertórios terminológicos e harmonizar os recursos obtidos com as entradas do VOC, é necessário um banco de dados desenhado especificamente para isso. Assim, a estrutura do banco de dados no qual os dados são alojados foi pensada de forma que: i) novas informações possam ser facilmente incorporadas; ii) o português seja tratado enquanto língua pluricêntrica (marcação de restrições geográficas); iii) haja a possibilidade de fácil conexão com o VOC; iv) e que permita a organização por domínio específico.

Por fim, para fortalecer a rede multinacional de especialistas em terminologias nos EMs da CPLP, criou-se a TECER como uma plataforma *web*, auxiliando, assim, a comunicação entre lexicógrafos.

4.1. Plataforma

Para cumprir todos os propósitos acima mencionados, precisávamos de ferramentas computacionais que permitissem a indexação e a validação dos termos. Também queríamos que fossem de código aberto e tivessem uma arquitetura modular para possibilitar uma expansão fácil.

Com esses requisitos em mente, criamos a TECER com um sistema integrado de indexação e validação on-line para o trabalho síncrono de todas as equipes nacionais. O TECER segue uma arquitetura cliente-servidor, escrito em PHP, JavaScript, MySQL, HTML e CSS. Essa configuração nos dá flexibilidade suficiente, facilitando a conexão ao VOC e possivelmente a outros bancos de dados. Por meio dessa arquitetura cliente-servidor, a maioria das operações ocorre no lado do servidor, em vez de ocorrer no computador dos usuários. Qualquer operação é, portanto, rápida, independentemente das especificações do computador do usuário, mantendo os dados e a interface atualizados. Para garantir que as páginas *web* sejam responsivas e intuitivas, utilizamos o Bootstrap, um *framework* de código-aberto que contém classes CSS básicas e pré-construídas, para o desenvolvimento *front-end*. Para garantir a escalabilidade e gerenciabilidade do TECER, adotamos uma arquitetura modular.

Ao projetar a TECER, tínhamos as duas seguintes suposições. Primeiro, buscamos flexibilidade e acreditamos que um mesmo modelo deve servir a vários propósitos. Nosso projeto está em constante desenvolvimento, então, o que é uma prioridade agora pode não ser no futuro. Além disso, a TECER não precisa ficar restrita ao projeto TCTC. Como a TECER possui código aberto, outras pessoas poderão utilizá-la, promovendo, dessa forma, o IILP.

Em segundo lugar, a facilidade de uso é de alta prioridade. Os usuários da TECER não são funcionários em tempo integral. Provavelmente são voluntários com pouco tempo disponível ou profissionais de diferentes áreas de atuação. Portanto, a interface deve ser fácil e atraente; o mais próximo possível de ser à prova de erros humanos; com uma curva de aprendizado muito pequena.

4.2. Os módulos

Projetos de lexicografia podem levar décadas para serem concluídos. O desenvolvimento de novas tecnologias, porém, é muito mais rápido. Manter os dados na menor unidade possível e em formato tabular garante que não importa como a tecnologia evolua, os recursos desenvolvidos ainda serão facilmente manipulados. Assim se evita a

discussão constante sobre qual é o formato ideal para dicionários, além de tornar a conexão com o VOC mais fácil.

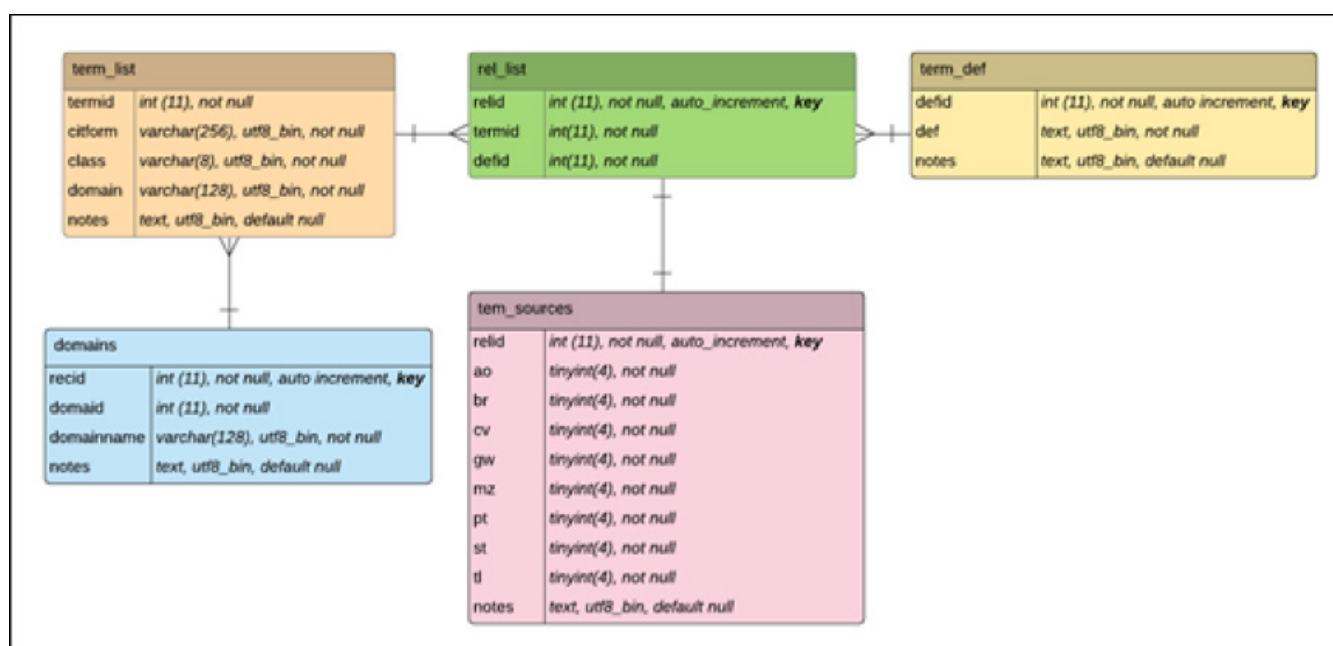
O mesmo princípio se aplica ao desenvolvimento de ferramentas para processar esses dados. Essas necessidades que temos agora podem (e provavelmente irão) mudar, e novas funções serão necessárias, enquanto outras se tornarão redundantes. Por isso, a TECER foi desenvolvida em módulos, e as novas funções são feitas de acordo com a necessidade. A TECER atualmente é composta por quatro módulos em constante desenvolvimento. Dois módulos básicos são (i) uma conexão com o VOC e (ii) um sistema de gerenciamento de usuários. Os dois módulos principais são (iii) uma ferramenta de validação e (iv) um sistema de indexação. Para estes dois últimos, presume-se que um determinado conjunto de dados já tenha sido pré-processado, mas ainda necessite inspeção humana.

4.2.1. Conexão com o VOC

A conexão com o VOC dá-se graças a estrutura do banco de dados (Figura 8), com as seguintes tabelas:

- term_list = lista de termos;
- term_def = definições ou conceitos, extraídos das terminologias de cada EM;
- rel_list = relações entre os termos e os conceitos;
- domain = indicação dos domínios, segundo o sistema de classificação que adotamos;
- term_sources = indicação da fonte/origem (EM).

FIGURA 8
Estrutura de banco de dados do TECER



Fonte: elaborado pelas autoras

Os termos listados em *term_list* podem ser facilmente associados aos termos listados no VOC, permitindo a criação de uma interface de conexão entre as terminologias do projeto TCTC e o VOC. É importante mencionar que há um passo intermediário (Figura 9) antes que os termos sejam inseridos na *term_list*.

FIGURA 9
Interface de conexão entre TECER e VOC

transport_domain01	
sourceid	int (11), not null, auto_increment, key
citform	varchar(256), utf8_bin, not null
class	varchar(8), utf8_bin, default null
termid	int (11), not null
def	text, utf8_bin, not null
equival	varchar (256), utf8_bin, default null
external_ref	varchar (128), utf8_bin, default null
notes	text, utf8_bin, default null

source_domain01	
recid	int (11), not null, auto increment, key
citform	varchar(256), utf8_bin, not null
def	text, utf8_bin, not null
notes	text, utf8_bin, default null
year	int (11), not null
autor	varchar(256), utf8_bin, not null
title	varchar(256), utf8_bin, not null

source_all	
table_name	varchar (128), utf8_bin, not null
domain	varchar (256), utf8_bin, not null
date_input	date, not null, default current_timestamp()
format_org	varchar (128), utf8_bin, not null
location	varchar (256), utf8_bin, not null
source	varchar (256), utf8_bin, not null
notes	text, utf8_bin, default null

Fonte: elaborado pelas autoras

O *tctc_sources* é um banco de dados com as fontes primárias em estado bruto. Isso acontece quando terminologias já processadas em formato tabular são obtidas. Quando isso ocorre, a inserção é feita diretamente na TECER (normalmente pela Equipe Central – EC), sem que o carregamento dos termos deva ser feito um por um.

Já o banco de dados *tctc_transport* é utilizado para fazer as transformações necessárias nas fontes primárias para que possam ser inseridas na base de dados mySQL e processadas, evitando, assim, que ocorram perdas ou alterações indevidas no banco de dados principal.

4.2.2. Gerenciamento de usuários

O sistema de gerenciamento de usuários é responsável por dois perfis: terminólogos e administradores, estes últimos capazes de conferir permissões, atribuir tarefas e monitorar o progresso de cada tarefa/usuário.

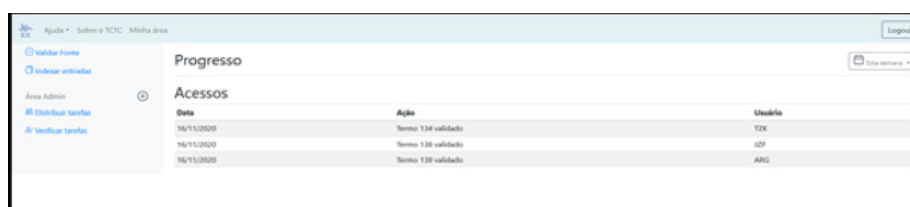
Cada membro das ENs terá um nome de usuário e uma senha para fazer *login* e acessar as terminologias referentes ao seu país. Os usuários com direitos de administrador (EC) poderão verificar o progresso do trabalho de cada colaborador. Para cada modificação no

banco de dados, são apresentados a data, a ação que foi efetuada (validação ou indexação) e o usuário (Figura 10). O sistema de gerenciamento de usuários é responsável por dois perfis: terminólogos e administradores, estes últimos capazes de conferir permissões, atribuir tarefas e monitorar o progresso de cada tarefa/usuário.

Cada membro das ENs terá um nome de usuário e uma senha para fazer *login* e acessar as terminologias referentes ao seu país. Os usuários com direitos de administrador (EC) poderão verificar o progresso do trabalho de cada colaborador. Para cada modificação no banco de dados, são apresentados a data, a ação que foi efetuada (validação ou indexação) e o usuário (Figura 10).

FIGURA 10

Tela de gerenciamento de usuários



Progresso		
Acessos		
Data	Ação	Usuário
16/11/2020	Termos 1.04 validado	TOK
16/11/2020	Termos 1.08 validado	IGP
16/11/2020	Termos 1.08 validado	ARG

Fonte: elaborado pelas autoras

4.2.3. Validação dos termos

A ferramenta de validação permite a verificação manual (validação e alteração da classificação anterior) ou automática de termos obtidos de várias fontes e inseridos automaticamente na plataforma. Este módulo também apresenta a classificação de aceitabilidade que os termos receberam de usuários de outros países de língua portuguesa que também possuem uma terminologia do mesmo domínio para sua variedade do idioma. A TECER adota a classificação de aceitabilidade estabelecida para o VOC por Janssen e Ferreira (2018); no entanto, o sistema pode ser facilmente adaptado para outros esquemas de classificação.

O protótipo apresentado na Figura 10 ilustra uma tarefa que envolve validar o termo encontrado na Classificação de Profissões de Angola. Evidentemente, no caso das profissões, em teoria, não seria necessário validar os termos, já que se trata de termos oficiais, elaborados por uma equipe especializada e usados amplamente nos países. Contudo, as ENs, se julgarem necessário, poderão fazer a validação desses termos. Uma razão para quererem fazer a validação poderia ser para revisar gralhas ortográficas, como se vê, por exemplo, no termo 2 (Oficial da Força Area, em vez de Oficial da Força Aérea). A EC poderia apresentar para as ENs apenas os termos que foram identificados, após análise automática, como contendo algum tipo de problema.

FIGURA 11
Tela de validação

fonte de citação	domínio (fonte)	domínio (indexador)	notas	validar
1 Oficial do Exército	prof.0110.1	Oficial do Exército		AO: 1 2 3 4 5 Ok
2 Oficial da Força Aérea	prof.0110.2	Oficial da Força Aérea		AO: 1 2 3 4 5 Ok
3 Oficial da Marinha	prof.0110.3	Oficial da Marinha		AO: 1 2 3 4 5 Ok
4 Sargento do Exército	prof.0210.1	Sargento do Exército		AO: 1 2 3 4 5 Ok

Fonte: elaborado pelas autoras

Como se vê na coluna mais à direita da Figura 10 (validar), para cada termo, o usuário é informado sobre a fonte do termo (nesse exemplo, AO = Angola) e a categoria de validação, que vai de 1 a 5. Inspirados na categorização adotada no projeto VOC, no TCTC essas categorias correspondem a:

1. termo explicitamente registrado em uma fonte primária para o país;
2. termo de uso irrestrito no país, mas não explicitamente registrado em uma fonte primária;
3. termo não recomendável no país devido a considerações de uso;
4. termo não representativo da terminologia do país devido às opções de variação do Acordo Ortográfico de 1990 (AO90) específicas do país;
5. termo totalmente inaceitável no país.

No caso das terminologias das profissões, como os termos foram extraídos de fontes primárias oficiais, o sistema automaticamente categoriza os termos como 1. Em terminologias de outros domínios poderá não ser o mesmo caso, como as do petróleo e da informática, quando estas forem transferidas para a TECER.

Além disso, nessa mesma coluna "validar", abaixo do número para cada categoria, é possível visualizar a que categoria aquele termo pertence em outras variedades do português. Para isso, basta passar o cursor por cima dos números em azul (Figura 12).

FIGURA 12

Visualizando a categoria de validação dos termos em outra(s) variedade(s) do português

validar

AO: 1 2 3 4 5 Ok

BR: 1 2 3 4 5

Fonte: elaborado pelas autoras

No caso usado aqui como exemplo, o fato de aparecer BR sobre o número 1 indica, para o usuário da EN angolana, que o mesmo termo foi também validado como categoria 1 pela EN brasileira. No futuro, poderá ocorrer que ENs cujos EMs não possuam uma terminologia para determinado domínio decidam adotar termos usados em outras variedades. Poderão, por exemplo, definir que assumirão que determinado termo passará a fazer parte da terminologia do seu país se este tiver sido categorizado como 1 em pelo menos um país, tendo em vista que, em se tratando de terminologia, é possível que somente um país detenha conhecimento técnico-científico em determinada área.

4.2.4. Indexação

A ferramenta de indexação permite o alinhamento dos termos de fontes múltiplas com um índice externo já existente (desejavelmente uma nomenclatura oficial do órgão regulador para um determinado domínio). Este módulo também permite aos usuários validar um código equivalente atribuído automaticamente do índice de referência a um termo na TECER ou fazê-lo manualmente.

No caso das classificações das profissões, como explicado acima, há um problema maior no processo de alinhamento da Classificação Brasileira das Ocupações (CBO) com a classificação dos outros países (Angola, Cabo Verde, Moçambique e Portugal), porque estes seguem a codificação adotada pela ISCO (*International Standard Classification of Occupations*), enquanto o Brasil criou uma codificação própria. Contudo, como a ISCO é em inglês, optamos por usar a ESCO (*European Skills, Competences and Occupations*), que é a classificação Europeia de profissões, competências e habilidades, criada também com base na ISCO, e que disponibiliza as classificações em 26 línguas, incluindo o português. Assim, a indexação de entradas para o domínio das profissões inclui uma etapa anterior, qual seja, a conversão para padronização dos códigos.

Como se pode observar na Figura 13, ao usuário da EN brasileira é solicitado verificar se a sugestão de termo (coluna do meio), retirado da ESCO se refere ao mesmo termo da CBO (primeira coluna). Se sim, o usuário clicará em ok na coluna converter (código) (última coluna) e ficará registrado, na base de dados, que o termo da CBO com o código X corresponde ao código Y da ESCO. Isso permitirá, depois, o alinhamento entre todas as variedades do português. Se o usuário tiver dúvidas, pode clicar no ponto de interrogação (?) ao lado dos termos para ler a sua definição, que aparecerá como *pop-up* (Figura 14).

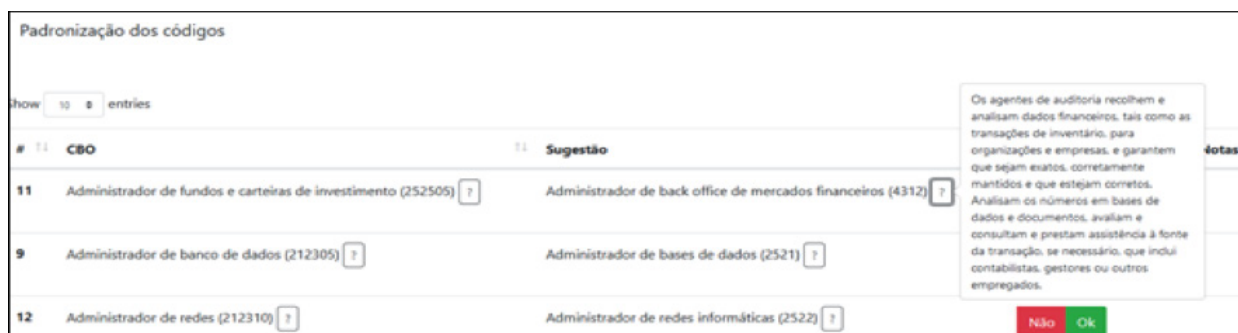
FIGURA 13

Tela de conversão para padronização dos códigos

#	CBO	Sugestão	Converter (código)	Notas
11	Administrador de fundos e carteiras de investimento (252505) ?	Administrador de back office de mercados financeiros (4312) ?	Não Ok	
9	Administrador de banco de dados (212305) ?	Administrador de bases de dados (2521) ?	Não Ok	
12	Administrador de redes (212310) ?	Administrador de redes informáticas (2522) ?	Não Ok	

Fonte: elaborado pelas autoras

FIGURA 14
Pop-up na tela de conversão



Fonte: elaborado pelas autoras

5. Tarefas da Equipe Central (EC) e das Equipes Nacionais (ENs)

Tal como para o VOC, o projeto TCTC foi desenvolvido por uma Equipe Central (EC) e por Equipes Nacionais (ENs), cada uma delas representando um Estado-membro (EM). Ambas tiveram tarefas específicas e fundamentais para o desenvolvimento pleno do projeto.

A EC, constituída por um grupo menor de pessoas, com formação técnica necessária para o bom curso do projeto, ocupou-se das seguintes tarefas:

1. coordenação geral de todo o trabalho;
2. produção de documentação para as ENs, isto é, redação de materiais que fossem necessários, especificando a metodologia de trabalho;
3. criação de um sistema de classificação baseado nos já existentes, alinhamento das áreas de conhecimento e seu agrupamento em grandes áreas;
4. oferecimento de cursos de formação em Terminologia às ENs, de modo a dar a conhecer a atividade terminológica como também padronizar a metodologia de trabalho;
5. identificação dos eventuais termos, de distintos campos do saber, já presentes no VOC, e atribuição de uma rubrica indicativa da área e/ou atividade em cada um deles;
6. submissão de listas parciais de termos (por área de conhecimento) às ENs para validação;
7. digitalização de repertórios terminológicos impressos e conversões de formatos para os que já estiverem em modo digital, obtidos por ENs cujos EMs encontraram dificuldade computacional de o fazer;
8. disponibilização das terminologias no VOC;
9. produção regular de relatórios para o IILP;
10. desenvolvimento de tarefas computacionais:
 - extração de terminologias de dicionários selecionados;
 - cotejamento dessas terminologias com as palavras do VOC;
 - criação de suporte para unidades polilexicais²⁰ no sistema OSLIN (sistema de bases de dados do VOC);
 - desenvolvimento da plataforma TECER.

²⁰ Termos constituídos por mais de uma palavra que requerem um tratamento linguístico e computacional especial.

Já as ENs, nomeadas pelos respectivos EMs, tiveram as seguintes atribuições:

1. identificação de especialistas de domínio que pudessem validar terminologias pertencentes às distintas áreas de conhecimento;
2. validação de listas de termos de áreas específicas enviadas pela EC;
3. obtenção de terminologias com garantia de fiabilidade²¹ já existentes no EM;
4. inventariação de terminologias e recursos relevantes existentes no EM;
5. inventariação de instituições representativas que dessem suporte ao projeto em nível nacional;
6. estabelecimento de contatos com organismos e/ou empresas que pudessem facilitar o trabalho terminológico;
7. e, futuramente, quando a plataforma estiver operante, inserção na TECER das terminologias obtidas ou envio em formato de planilha Excel para a EC realizar a inserção.

É importante destacar que o sucesso do projeto TCTC dependeu do engajamento dos EMs, pois sem o trabalho ativo das ENs, as terminologias em língua portuguesa perdem em visibilidade, pois não constarão do VOC.

6. Considerações Finais

Procuramos trazer ao conhecimento do público mais amplo as origens do projeto TCTC e a sua relevância no cenário da CPLP.

Apresentamos os princípios teóricos norteadores, nos quais se assentou o referido projeto, como também detalhamos as etapas metodológicas que guiaram o trabalho das ENs, tanto para o alinhamento de termos entre as variedades do português quanto para a elaboração de uma terminologia de raiz.

Expusemos a plataforma computacional TECER, criada para atender os objetivos do projeto TCTC, facilitando enormemente o trabalho das ENs e da EC, mas que também poderá servir para outros projetos e línguas, atendendo, assim, a diferentes propósitos e promovendo o IILP.

Por fim, demos a conhecer as tarefas que ficaram a cargo da EC e das ENs, enaltecendo a importância destas últimas, sem as quais não há maneira de sistematizar as terminologias em língua portuguesa.

7. Referências Bibliográficas

ALMEIDA, G. M. B. *Teoria Comunicativa da Terminologia: uma aplicação*. Araraquara, vol. 1, p. 290; vol. 2, p.86. Tese (Doutorado em Linguística e Língua Portuguesa) – Faculdade de Ciências e Letras, Câmpus de Araraquara, Universidade Estadual Paulista, 2000.

²¹ Entende-se que uma terminologia tenha um bom índice de fiabilidade quando se comprova que os seus termos constitutivos são efetivamente usados pelos especialistas do domínio científico ou técnico em questão, ou quando resultantes de uma proposta do terminólogo, sejam bem construídos de acordo com as regras linguísticas e aceites pelos especialistas desse domínio.

ALMEIDA, G. M. B. *A Teoria Comunicativa da Terminologia e a sua prática*. Alfa (Araraquara), v. 50, 2006, pp. 81-97. Disponível em: <http://www.alfa.ibilce.unesp.br/download/v50-2/06-Almeida.pdf>. Acesso em: 23 set. 2009.

ALMEIDA, G. M. B. *Fazer Terminologia é fazer Linguística*, in: PERNA, C. L.; DELGADO, E. K.; FINATTO, M. J. B. (orgs.) *Linguagens especializadas em corpora: modos de dizer e interfaces de pesquisa [recurso eletrônico]* Porto Alegre: EDIPUCRS, 2010, p. 255. Disponível em: <http://www.pucrs.br/edipucrs/linguagensespecializadasemcorpora.pdf>. Acesso em: 12 out. 2010.

ALMEIDA, G. M. B.; OLIVEIRA, L. H. M.; ALUISIO, S. M. *A terminologia na era da informática*. Cienc. Cult. [online]. vol. 58, n° 2 [citado 2010-03-14], 2006, pp. 42-45. Disponível em: http://cienciaecultura.bvs.br/scielo.php?script=sci_arttext&pid=S0009-67252006000200016. Acesso em: 3 nov. 2021.

ALMEIDA, G. M. B.; ALUÍSIO, S. M.; OLIVEIRA, L. H. M. *O método em Terminologia: revendo alguns procedimentos*. In: Aparecida Negri Isquerdo; Ieda Maria Alves (orgs.). *Ciências do léxico: lexicologia, lexicografia, terminologia*. 1ª ed. Campo Grande/São Paulo: Editora da UFMS/Humanitas, v. 3, 2007, pp. 409-420.

ALMEIDA, G. M. B.; CORREIA, M. *Terminologia e corpus: relações, métodos e recursos*, in: Stella E. O. Tagnin; Oto Araújo Vale. (orgs.). *Avanços da Linguística de Corpus no Brasil*. 1ª ed. São Paulo: Humanitas/FFLCH/USP, v. 1, 2008, pp. 67-94.

ALUÍSIO, S. M.; ALMEIDA, G. M. B. *O que é e como se constrói um corpus? Lições aprendidas na compilação de vários corpora para pesquisa linguística*. Calidoscópio (UNISINOS). Vol. 4, n. 3, 2006, pp. 155-177, set/dez. Disponível em: http://www.unisinis.br/publicacoes_cientificas/images/stories/pdfs_calidoscopio/vol4n3/art04_aluisio.pdf. Acesso em 12 out. 2010.

ALMEIDA, G. M. B. et.al. *Glossário de revestimento cerâmico*. In: Cadernos de terminologia, n° 04, 2011. pp. 03-56. Universidade de São Paulo (USP) – Faculdade de Filosofia, Letras e Ciências Humanas: Centro Interdepartamental de Tradução e Terminologia (CITRAT), 2011. Disponível em <https://citrat.fflch.usp.br/sites/citrat.fflch.usp.br/files/inline-files/Cad.%20Terminologia%204.pdf>. Acesso em: 8 nov. 2021.

ALMEIDA, G. M. B.; OLIVEIRA, L.H.M. *Ambiente web para elaboração de produtos terminológicos: e-Termos*. In: Éric Laporte, Aucione Smarsaro, Oto Araújo Vale (orgs.). *Dialogar é preciso: linguística para processamento de línguas*. 1ed. Vitória: PPGEL/UFES, 2013, v. 1, p. 21-30. Disponível em: https://www.researchgate.net/publication/280883367_Ambiente_web_para_elaboracao_de_produtos_terminologicos_e-Termos. Acesso em: nov. 2021.

BARROS, L. A. *Curso básico de terminologia*. São Paulo: EDUSP, 2004.

BERBER SARDINHA, T. *Linguística de corpus*. São Paulo: Manole, 2004.

BIBER, D.; CONRAD, S.; REPPEN, R. *Corpus linguistics: investigating language structure and use*. Cambridge University Press, Cambridge, 1998.

BOULANGER, J. C. *Présentation: images et parcours de la socioterminologie*. Meta, vol. 40, nº 2, 1995.

CABRÉ, M. T. *La terminología – teoría, metodología, aplicaciones* (trad. castelhana de Carles Tebé). Barcelona: Editorial Antártida/Empúries, 1993.

CABRÉ, M. T. (org.). *Terminologia – selección de textos d'E. Wüster* (trad. catalã de Anna C. i Galera e Teresa G. i Morell). Barcelona: Servei de Llengua Catalana – Universitat de Barcelona, 1996.

CABRÉ, M. T. *La terminología: representación y comunicación – elementos para una teoría de base comunicativa y otros artículos*. Barcelona: Institut Universitari de Lingüística Aplicada, 1999.

CABRÉ, M. T. *Theories of Terminology: their description, prescription and explanation*. Terminology, vol. 9, nº 2, 2003, pp. 163-200.

CIAPUSCIO, G. E. *Textos especializados y terminología. Sèrie Monografies*. Barcelona: IULA/ Universitat Pompeu Fabra, 2003.

CPLP – COMUNIDADE DE PAÍSES DE LÍNGUA PORTUGUESA, *Plano de ação de Brasília – PAB*. Brasília, 2010.

CPLP – COMUNIDADE DE PAÍSES DE LÍNGUA PORTUGUESA - CPLP, *Plano de ação de Lisboa*. Lisboa, 2013. – PALis. disponível em: https://www.cplp.org/Files/Billeder/cplp/1-Resol_PALis.pdf. Acesso em: 21 nov 2021.

FELBER, H. *En memòria d'E. Wüster*. In: CABRÉ, M. T. *Terminologia – selecció de textos d'E. Wüster* (trad. catalã de Anna C. i Galera e Teresa G. i Morell). Barcelona: Servei de Llengua Catalana – Universitat de Barcelona, 1996, pp.17-29.

FERREIRA, J. P.; ALMEIDA, G.M.B. *Manual para elaboração de corpora - com vista à organização dos Vocabulários Ortográficos Nacionais dos países integrantes da CPLP*. Praia: Instituto Internacional da Língua Portuguesa (IILP), 2012.

GALINSKY, C. *La figura d'E. Wüster*. In: CABRÉ, M. T. *Terminologia – selección de textos d'E. Wüster* (trad. catalã de Anna C. i Galera e Teresa G. i Morell). Barcelona: Servei de Llengua Catalana – Universitat de Barcelona, 1996, pp.11-15.

GAMBIER, Y. *Problèmes terminologiques des pluies acides*. Meta, vol. 32, nº 3, 1987.

GAMBIER, Y. *Travail et vocabulaire spécialisés: prolégomènes à une socio-terminologie*. Meta, vol. 36, nº 1, 1991.

GAUDIN, F. *Socioterminologie. Des problèmes semantiques aux pratiques institutionnelles*. Rouen: Université de Rouen, 1993.

GEWEHR, D.; NEIDE, I. G.; DULLIUS, M. M. *Mapas conceituais com CmapTools: uma metodologia ativa de ensino e aprendizagem para nativos digitais*. Educitec, Manaus, v. 04, n. 07, p. 152-165, jun. 2018.

HOFFMAN, L. *Els llenguatges d'especialitat: selecció de textos. Sèrie Monografies*. Barcelona: IULA/Universitat Pompeu Fabra, 1998

JANSSEN, M.; FERREIRA, J. P. *Technical Implementation of the Vocabulário Ortográfico Comum da Língua Portuguesa*. In: Villavicencio, Alina (eds.) *Computational Processing of the Portuguese Language*. PROPOR 2018, Cham: Springer, 2018.

KAMIKAWACHI, D. S. L. *Aspectos semânticos da definição terminológica (DT): descrição linguística e proposta de sistematização*. Dissertação (Mestrado em Linguística) – Campus de São Carlos, Universidade Federal de São Carlos, 2009. Disponível em http://200.136.241.56/htdocs/tedeSimplificado/tde_busca/arquivo.php?codArquivo=3033&PHPSESSID=e2985a9c2d5a6b1420ff29ea56f9722a. Acesso em: 12 out. 2010.

KENNEDY, G. *An Introduction to Corpus Linguistics*, London/NY: Longman, 1998.

KRIEGER, M. G. *Terminologia revisitada*. DELTA: Documentação e Estudos em Linguística Teórica e Aplicada, São Paulo, v. 16, n. 2, 2000

KRIEGER, M. G.; BEVILACQUA, C. R. *A pesquisa terminológica no Brasil: uma contribuição para a consolidação da área*. Debate Terminológico, no 1, 03, 2004. Disponível em: <http://riterm.net/revista/ojs/index.php/debateterminologico/article/view/24>. Acesso em 12 out. 2010.

KUHN, T. Z.; FERREIRA, J. P. *O corpus de português escrito em periódicos – CoPEP*. DELTA: Documentação e Estudos em Linguística Teórica e Aplicada, 36(2), 2020. Disponível em: <https://revistas.pucsp.br/index.php/delta/article/view/52552>. Acesso em nov 2021.

L'HOMME, M. C. *La terminologie: principes et techniques*. Montréal: Presses de l'Université de Montréal, 2004

LERAT, P. *Las lenguas especializadas* (trad. castelhana de Albert Ribas). Barcelona: Editorial Ariel, 1997.

LORENTE, M. *A lexicologia como ponto de encontro entre a gramática e a semântica*. In: ISQUERDO, A. N. e KRIEGER, M.G. (orgs.) *As ciências do léxico*, vol. 2. Campo Grande: Editora UFMS, 2004.

MUNIZ, M. C. M. *A construção de recursos linguístico-computacionais para o português do Brasil: o projeto de Unitex-PB*. Dissertação de Mestrado. Instituto de Ciências Matemáticas de São Carlos, USP, 2004.

PAUMIER, S. *Unitex user manual*, 2002. Disponível em: <http://www-igm.univ-mlv.fr/~unitex>. Acesso em: 18 jun. 2008.

PEARSON, J. *Terms in Context*. Amsterdam / Philadelphia: John Benjamins Publishing Company, 1998.

RENOUF, A. *Explorations in Corpus Linguistics*. Amsterdam: Rodopi, 1998.

SAGER, J. C. *Curso práctico sobre el procesamiento de la terminología*. (trad. castelhana de Laura C. Moya). Madrid: Fundación Germán Sánchez Ruipérez/Pirámide, 1993.

SINCLAIR, J. *Corpus and Text - Basic Principles*. In: *Developing Linguistic Corpora: a Guide to good Practice*, ed. M. Wynne. Oxford: Oxbow Books, 2005, pp.1-16. Disponível em: <http://ahds.ac.uk/linguistic-corpora/>. Acesso em: 18 jun. 2008.

TEMMERMANN, R. *Towards new Ways of Terminology Description. The Sociocognitive Approach*. Philadelphia: John Benjamins, 2000, p. 258.

SOUZA, J. A. *Tipologia de traços linguísticos de textos do português do Brasil dos séculos XVI, XVII, XVIII e XIX: uma proposta para a classificação automática de gêneros textuais*. Dissertação (Mestrado). São Carlos: UFSCar, 2011. Disponível em: <https://repositorio.ufscar.br/handle/ufscar/5698>. Acesso em: 6 nov. 2021.

SUHARDIJANTO, T.; DINAKARAMANI, A. *Building a Collaborative Workspace for Lexicography Works in Indonesia*. The Fifth ELex Conference Electronic Lexicography in the 21st Century, 299–308, 2017.

8. Apêndice 1 - Domínios discursivos e tipos textuais²²

1. DOMÍNIO DISCURSIVO – CIENTÍFICO

1.1 Divulgação

- 1.1.1 Artigo
- 1.1.2 Ensaio
- 1.1.3 Resenha
- 1.1.4 Resumo

1.2 Pesquisa

- 1.2.1 Artigo
- 1.2.2 Dissertação
- 1.2.3 Monografia
- 1.2.4 Projeto
- 1.2.5 Tese

²² Com base em Souza (2011).

2 DOMÍNIO DISCURSIVO – INSTRUCIONAL

2.1 Didático

- 2.1.1 Apostila
- 2.1.2 Livro-texto

2.2 Procedimental

- 2.2.1 Bula
- 2.2.2 Manual
- 2.2.3 Receita

3 DOMÍNIO DISCURSIVO – INFORMATIVO

3.1 Jornalístico

- 3.1.1 Artigo
- 3.1.2 Editorial
- 3.1.3 Entrevista
- 3.1.4 Reportagem

3.2 Informe

- 3.2.1 Aviso
- 3.2.2 Boletim
- 3.2.3 Comunicado

4 DOMÍNIO DISCURSIVO – TÉCNICO-ADMINISTRATIVO (ESFERA PÚBLICA OU PRIVADA)

4.1 Comunicacional

- 4.1.1 Ato
- 4.1.2 Carta
- 4.1.3 Circular
- 4.1.4 Declaração
- 4.1.5 Despacho
- 4.1.6 Memorando
- 4.1.7 Ofício
- 4.1.8 Provisão
- 4.1.9 Requerimento
- 4.1.10 Solicitação

4.2 Descritivo

- 4.2.2 Ata
- 4.2.3 Auto de exame médico
- 4.2.4 Balanço (financeiro)
- 4.2.5 Informe
- 4.2.6 Levantamento
- 4.2.7 Ordem do dia

4.2.8 Planejamento

4.2.9 Regimento

4.2.10 Relatório

4.2.11 Memorial

4.3 Comercial

4.3.1 Conhecimento

4.3.2 Contrato

4.3.3 Nota

4.3.4 Recibo

5 DOMÍNIO DISCURSIVO – REFERENCIAL

5.1 Catálogo

5.2 Dicionário

5.3 Glossário

5.4 Índice

5.5 Verbete

5.6 Thesaurus

6 DOMÍNIO DISCURSIVO – JURÍDICO

6.1 Legislativo

6.1.1 Códigos

6.1.2 Constituição

6.1.3 Decreto

6.1.4 Decreto lei

6.1.5 Emenda constitucional

6.1.6 Lei

6.1.7 Medida provisória

6.1.8 Portaria

6.1.9 Resolução

6.1.10 Tratado internacional

6.2 Jurisprudência

6.2.1 Acórdão

6.2.2 Minuta

6.2.3 Petição

6.2.4 Sentença

6.2.5 Súmula

6.3 Jurídico-administrativo

6.3.1 Atos administrativos

6.3.1.1 Alvará

6.3.1.2 Autorização

6.3.1.3 Licença

6.3.1.4 Admissão

6.3.1.5 Permissão

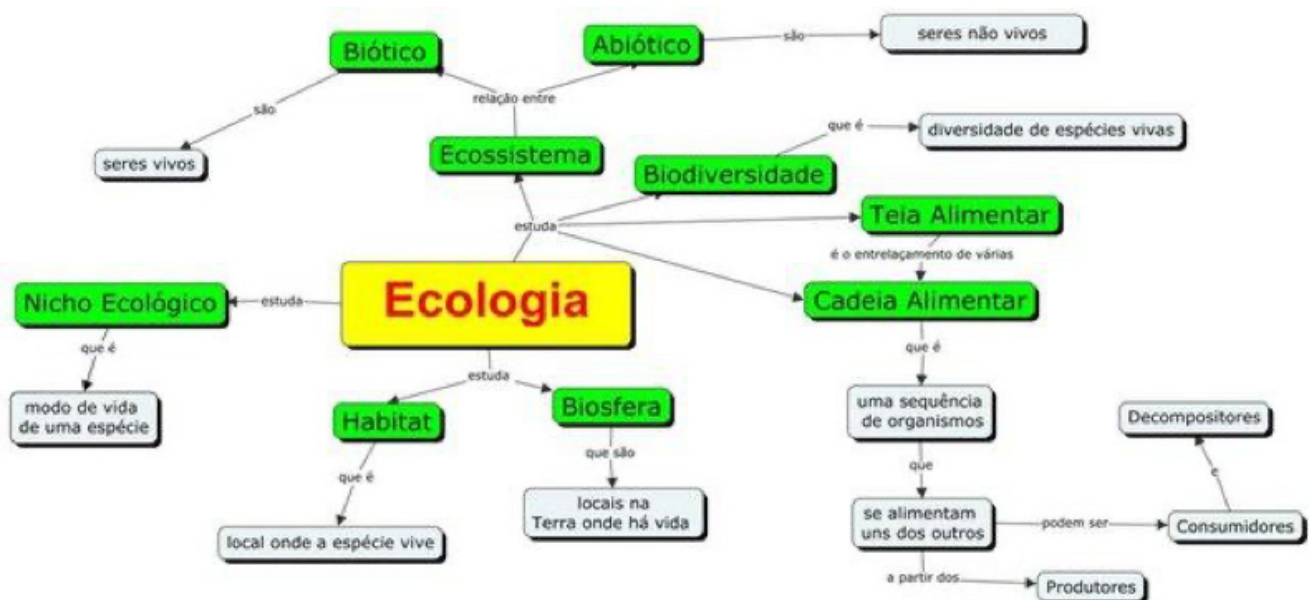
- 6.3.1.6 Aprovação
- 6.3.1.7 Homologação
- 6.3.1.8 Parecer
- 6.3.1.9 Visto
- 6.3.1.10 Exposição de motivos
- 6.3.1.11 Decreto
- 6.3.1.12 Resolução
- 6.3.1.13 Portaria
- 6.3.1.14 Circular
- 6.3.1.15 Depoimento
- 6.3.2 Licitação
 - 6.3.2.1 Edital
 - 6.3.2.2 Concorrência
 - 6.3.2.3 Convite
 - 6.3.2.4 Pregão
 - 6.3.2.5 Tomada de preços
- 6.3.3 Contrato
 - 6.3.3.1 Concessão
 - 6.3.2.2 Contrato de obra pública e de prestação de serviço
 - 6.3.2.3 De fornecimento
 - 6.3.2.4 De gestão
 - 6.3.2.5 De convênio
 - 6.3.2.6 De consórcio
- 6.3.4 Administrativo Público Notarial
 - 6.3.4.1 Testamento
 - 6.3.4.2 Devassa
 - 6.3.4.3 Inventário
 - 6.3.4.4 Certidão
 - 6.3.4.5 Certificado
 - 6.3.4.6 Trespasse
 - 6.3.4.7 Sumário de testemunha
 - 6.3.4.8 Procuração
- 6.3.5 Abaixo-assinado
- 6.3.6 Ata
- 6.3.7 Atestação
- 6.3.8 Atestado
- 6.3.9 Auto
- 6.3.10 Foral
- 6.3.11 Lançamento
- 6.3.12 Notificação
- 6.3.13 Ofício
- 6.3.14 Carta patente
- 6.3.15 Carta precatória
- 6.3.16 Provimento
- 6.3.17 Termo
 - 6.3.17.1 Termo de declaração
 - 6.3.17.2 Termo de vereação

- 6.3.17.3 Termo de junta
- 6.3.17.4 Termo de testemunha

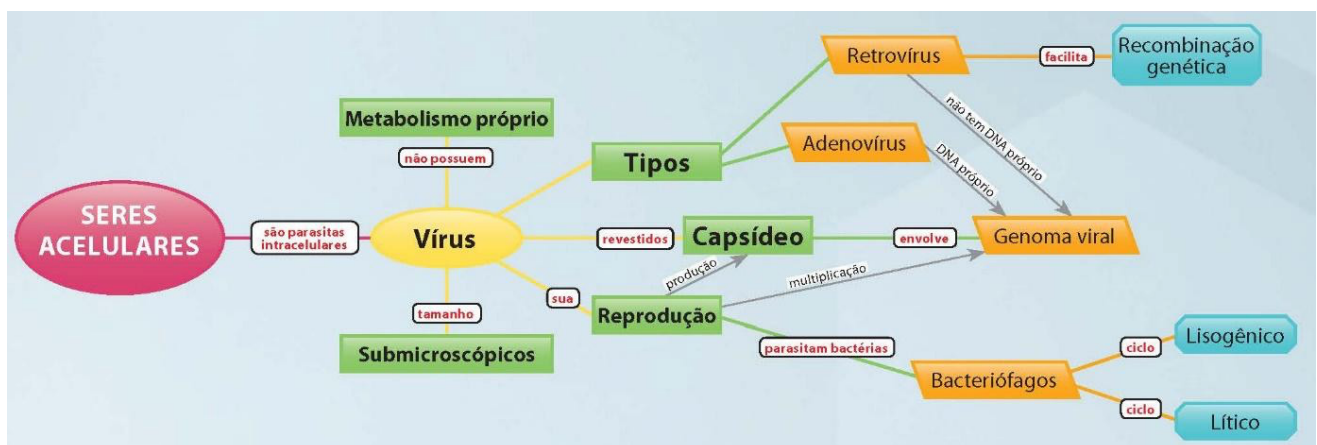
6.4 Jurídico Comercial

- 6.4.1 Escritura
- 6.4.2 Registro
- 6.4.3 Representação

9. Anexo 1 - Exemplos de estruturas conceituais elaboradas no CMAPTools



Fonte: GEWEHR et al. (2018)



Fonte: <<https://educacional.cpb.com.br/conteudos/conhecimento-conteudos/mapas-conceituais-em-educacao-cientifica/>>. Acesso em 21 nov. 2021

Apresentação – Instituto Internacional da Língua Portuguesa

Apresentação – Ministério das Relações Exteriores

Apresentação – Coordenação do Projeto Terminologias Científicas
e Técnicas Comuns da Língua Portuguesa

Projetos pluricêntricos do IILP. Apontamentos para a história da política
linguística do português

Variedades nacionais de línguas pluricêntricas: o caso do português em
Moçambique

A importância das terminologias em português para a tradução e a
interpretação na União Europeia

A língua portuguesa no espaço dos países africanos de língua oficial
portuguesa - da variedade do português de São Tomé e Príncipe à
definição de um padrão

A língua portuguesa no espaço dos países africanos de língua oficial
portuguesa – uma reflexão sobre Cabo Verde

Construindo uma terminologia de raiz: textos legislativos sob exploração
terminológica

Língua e diplomacia – uma agenda para o diálogo em língua portuguesa

Procedimentos teórico-metodológicos para recolha e sistematização de
terminologias no âmbito da Comunidade de Países de Língua Portuguesa
(CPLP)

