



Departamento de Sociologia

## Topónimos no espaço da CPLP: o Vocabulário Toponímico

Agostinho Miguel Magalhães Salgueiro

Dissertação submetida como requisito parcial para obtenção do grau de  
Mestre em Comunicação, Cultura e Tecnologias da Informação

Orientador:

Doutor José Soares da Silva Neves, Professor Auxiliar Convidado  
ISCTE – Instituto Universitário de Lisboa

Coorientadora:

Doutora Margarita Maria Correia Ferreira, Professora Auxiliar  
Faculdade de Letras – Universidade de Lisboa

Setembro, 2016





Departamento de Sociologia

## Topónimos no espaço da CPLP: o Vocabulário Toponímico

Agostinho Miguel Magalhães Salgueiro

Dissertação submetida como requisito parcial para obtenção do grau de  
Mestre em Comunicação, Cultura e Tecnologias da Informação

Orientador:

Doutor José Soares da Silva Neves, Professor Auxiliar Convidado  
ISCTE – Instituto Universitário de Lisboa

Coorientadora:

Doutora Margarita Maria Correia Ferreira, Professora Auxiliar  
Faculdade de Letras – Universidade de Lisboa

Setembro, 2016



## Agradecimentos:

Um muito sentido agradecimento aos meus queridos pais, Maria Idalina Magalhães Salgueiro e Agostinho Francisco Vieira Salgueiro.

Obrigado a todos quantos também contribuíram, em alguma medida, para a feitura desta dissertação.

## Resumo:

A toponímia é o ramo da onomástica que estuda os nomes atribuídos a lugares. De entre as abordagens que a disciplina explora, importa neste trabalho aquela que está diretamente relacionada com repositórios toponímicos padronizados: bases de dados lexicais toponímicos. A adequada padronização em língua portuguesa da toponímia maior no espaço da CPLP assume especial relevância, pela função normalizadora que desempenha e pela abrangência que a caracteriza. O Vocabulário Toponímico (VT), aqui estudado, é um recurso toponímico digital, um vocabulário de especialidade que é parte destacada do Vocabulário Ortográfico Comum da Língua Portuguesa (VOC). Com aproximadamente 70000 vocábulos padronizados, o VT inclui todos os topónimos com relevância administrativa no espaço da CPLP – quase exclusivamente endónimos –, bem como os nomes de todos os países e das suas capitais – maioritariamente exónimos.

Nesta dissertação foram analisados os métodos seguidos para a constituição da nomenclatura do vocabulário, as implicações da partição entre toponímia maior e toponímia menor e a relevância sociolinguística da toponímia, em geral, e de um recurso como o VT, em particular. Examinou-se cada subconjunto de topónimos, tendo em conta as relações hierárquicas estabelecidas entre as realidades que denominam, a sua distribuição geográfica e as relações derivacionais entre topónimos e gentílicos.

Concluimos que a padronização da toponímia e a possibilidade de verificação sistemática de relações toponímicas são de importância maior para a descrição de uma língua viva, permitindo referenciação geográfica inequívoca e homogénea e consolidando papéis identitários.

Palavras-chave: topónimos; padronização toponímica; Vocabulário Toponímico; Vocabulário Ortográfico Comum da Língua Portuguesa; CPLP.

## Abstract:

Toponymy is the branch of Onomastics that studies place names. From among the subfields in Toponymy, in this work we focus on standardized toponymic repositories: toponymic lexical data bases. An adequate standardization of the major toponymy in the CPLP area assumes special relevance, as it embodies a normalizing function and is characterized by its enlarged scope. The Vocabulário Toponímico (VT – toponymic wordlist), studied here, is a digital toponymic resource, a specialized vocabulary that is a relevant part of the Vocabulário Ortográfico Comum da Língua Portuguesa (VOC – the common Portuguese language wordlist). With nearly 70000 standardized vocables, the VT includes every toponym corresponding to an administrative division in the participating CPLP countries – almost exclusively endonyms –, as well as the name for every country and capital in the world – a large part of which are exonyms.

In this dissertation we analyzed the methodology used to gather the main entries of the wordlist, as well as the sociolinguistic relevance of toponymy, in general, and of a resource like the VT, in particular. We examined each subset of toponyms, taking into account the hierarchical relations between the realities they denote, their geographic distribution and the derivational relations between toponyms and gentiles.

We concluded that the standardization of toponyms and the possibility of systematic verification of toponymic relations are of major importance to the description of a living language, supporting unambiguous geographical references and the consolidation of identity roles.

Key-words: toponyms; toponymic standardization; Vocabulário Toponímico; Vocabulário Ortográfico Comum da Língua Portuguesa; CPLP.





# ÍNDICE

<u>Índice.....</u>	<u>ix</u>
<u>Lista de Quadros.....</u>	<u>x</u>
<u>Lista de Figuras.....</u>	<u>x</u>
<u>Glossário de siglas e acrónimos.....</u>	<u>xi</u>
<u>1. Introdução.....</u>	<u>13</u>
<u>2. Toponímia.....</u>	<u>15</u>
<u>2.1. Toponímia: perspetiva sociolinguística.....</u>	<u>16</u>
<u>2.2. Outras questões com implicações sociológicas.....</u>	<u>20</u>
<u>2.3. Outros dados linguísticos.....</u>	<u>23</u>
<u>3. Normalização linguística e recursos gerais.....</u>	<u>29</u>
<u>3.1. Norma ortográfica e variação lexical.....</u>	<u>29</u>
<u>3.2. VOC: enquadramento e papel normalizador.....</u>	<u>32</u>
<u>4. O Vocabulário Toponímico.....</u>	<u>35</u>
<u>4.1. Nomeação do recurso.....</u>	<u>35</u>
<u>4.2. Padronização: mais-valias.....</u>	<u>36</u>
<u>4.3. Critérios lexicais: seleção de topónimos.....</u>	<u>37</u>
<u>4.4. Filiação histórica na área da filologia.....</u>	<u>40</u>
<u>4.5. Tradição normativa para a toponímia.....</u>	<u>40</u>
<u>4.6. Toponímia em língua portuguesa e o UNGEGN.....</u>	<u>46</u>
<u>5. Dados do Vocabulário Toponímico.....</u>	<u>49</u>
<u>5.1. Países e capitais.....</u>	<u>49</u>
<u>5.2. Divisões administrativas da CPLP (e dependências).....</u>	<u>56</u>
<u>5.3. Uniformização e relações entre unidades toponímicas.....</u>	<u>60</u>
<u>5.4. Codificação.....</u>	<u>60</u>
<u>6. Conclusões e trabalho futuro.....</u>	<u>63</u>
<u>6.1. Abrangência e utilidade do recurso.....</u>	<u>63</u>
<u>6.2. Desenvolvimentos e aplicações futuros.....</u>	<u>64</u>
<u>Bibliografia.....</u>	<u>67</u>
<u>Webografia.....</u>	<u>69</u>
<u>Anexos.....</u>	<u>I</u>

## Lista de Quadros

- Quadro 5.1: Dupla grafia em nomes de países – casos com vogal tónica seguida de consoante nasal
- Quadro 5.2: Dupla grafia em nomes de países – casos com motivação política e ideológica
- Quadro 5.3: Dupla grafia em nomes de países – casos com terminação nasal
- Quadro 5.4: Dupla grafia em nomes de países – caso decorrente da pluralização do topónimo
- Quadro 5.5: Dupla grafia em nomes de países – adaptação díspar
- Quadro 5.6: Dupla grafia em nomes de países – variação decorrente de produção oral díspar
- Quadro 5.7: Dupla grafia em nomes de países – diferente tonicidade
- Quadro 5.8: Dupla grafia em nomes de países – variação puramente ortográfica
- Quadro 5.9: Dupla grafia em nomes de capitais – diferente tonicidade
- Quadro 5.10: Dupla grafia em nomes de capitais – casos com terminação nasal
- Quadro 5.11: Dupla grafia em nomes de capitais – casos com abertura de vogal final díspar
- Quadro 5.12: Dupla grafia em nomes de capitais – variação decorrente de produção oral díspar
- Quadro 5.13: Dupla grafia em nomes de capitais – variação puramente ortográfica
- Quadro 5.14: Nomes de capitais de países não adaptados ortograficamente
- Quadro 5.15: Gentílicos decorrentes de topónimos paralelos
- Quadro 5.16: Sufixo derivacional distinto empregue tendo por base o mesmo topónimo
- Quadro 5.17: Estado-unidense *versus* norte-americano
- Quadro 5.18: Gentílicos derivados de topónimos distintos com um só referente geográfico (Myanmar/Birmânia)
- Quadro 5.19: Pares de gentílicos comuns a todas as variedades e derivados de um único topónimo
- Quadro 5.20: Subdivisões do Brasil (administrativas e não administrativas)
- Quadro 5.21: Subdivisões de Cabo Verde (administrativas e não administrativas)
- Quadro 5.22: Subdivisões de Moçambique (administrativas e não administrativas)
- Quadro 5.23: Subdivisões de Portugal (administrativas e não administrativas)
- Quadro 5.24: Subdivisões administrativas de São Tomé e Príncipe
- Quadro 5.25: Subdivisões administrativas de Timor-Leste
- Quadro 5.26: Subdivisões administrativas de Angola
- Quadro 5.27: Subdivisões administrativas da Guiné Bissau
- Quadro A: Países com linhas orientadoras toponímicas para uso internacional submetidas à UNGEGN
- Quadro B: Países reconhecidos pela ONU (janeiro de 2016): grafia normalizada pelo VT

## Lista de Figuras

- Figura A: Placa toponímica de Armação de Pera, agosto de 2009
- Figura B: Placa toponímica para Armação de Pera, agosto de 2009
- Figura C: Placa toponímica de Troia, agosto de 2009
- Figura D: Placa toponímica de Troia, dezembro de 2014
- Figura E: Placa toponímica para Troia, agosto de 2009
- Figura F: Placa toponímica para Troia, janeiro de 2015
- Figura G: Placas toponímicas de Pero Pinheiro, setembro de 2009
- Figura H: Placas toponímicas de Pero Pinheiro, setembro de 2014

Figura I: Placa toponímica de Crastos, novembro 2009

Figura J: Placa toponímica de Crastos, fevereiro de 2015

## **Glossário de siglas e acrónimos**

AO45: Acordo Ortográfico da Língua Portuguesa, de 1945

AO90: Acordo Ortográfico da Língua Portuguesa, de 1990

CAOP: Carta Administrativa Oficial de Portugal

CELGA: Centro de Estudos de Linguística Geral e Aplicada

CELGA-ILTEC: Centro de investigação da Universidade de Coimbra (fusão entre o CELGA e o ILTEC)

CIA: Central Intelligence Agency

CPLP: Comunidade dos Países de Língua Portuguesa

DFNA: Dicionário de Formas Não Adaptadas

FO11: Formulário Ortográfico de 1911

FO43: Formulário Ortográfico de 1943

GPNUNG: Grupo de Peritos das Nações Unidas em Nomes Geográficos (ver UNGEGN)

IBGE: Instituto Brasileiro de Geografia e Estatística

IILP: Instituto Internacional da Língua Portuguesa

ILTEC: Instituto de Linguística Teórica e Computacional

INE: Instituto Nacional de Estatística

ISO: International Organization for Standardization

NILC: Núcleo Interinstitucional de Linguística Computacional

NUTS: Nomenclatura das Unidades Territoriais para fins Estatísticos

ONU: Organização das Nações Unidas

OSLIN: Open Source Lexical Information Network

TOP: Topónimo

UNGEGN: United Nations Group of Experts on Geographical Names

URSS: União das Repúblicas Socialistas Soviéticas

VOC: Vocabulário Ortográfico Comum da Língua Portuguesa

VOCALP: Vocabulário Cabo-Verdiano da Língua Portuguesa

VOLP: Vocabulário Ortográfico da Língua Portuguesa – Brasil

VOMOLP: Vocabulário Moçambicano da Língua Portuguesa

VOP: Vocabulário Ortográfico do Português – Portugal

VO-TL: Vocabulário Ortográfico de Timor-Leste

VT: Vocabulário Toponímico



# 1. INTRODUÇÃO

O nome de um lugar e a forma gráfica empregue para o representar são veículos de importância maior na identificação não ambígua de um referente geográfico por parte dos utilizadores de qualquer língua. Numa sociedade em rede<sup>1</sup> e do conhecimento (Ávila, 2008:16-18), torna-se fundamental ter à disposição uma obra normalizadora de referência para a língua portuguesa que ateste, pelo menos, os topónimos com relevância administrativa no espaço da Comunidade dos Países de Língua Portuguesa (CPLP), sob uma grafia comum, atual, com possibilidade de verificação das diferenças entre cada variedade nacional, sejam elas previsíveis ou não, e acessível a todos os falantes de forma gratuita e transparente, sem limitações quanto à informação fundamental passível de ser partilhada, incluindo propriedades formais. O Vocabulário Toponímico (doravante VT) assume esse papel normalizador para a toponímia em língua portuguesa, seguindo necessariamente as regras prescritas pelo Acordo Ortográfico da Língua Portuguesa, de 1990 (AO90).

A dissertação que se lê partilha informação diversificada relativa ao processo de elaboração e aos dados do VT – recurso da plataforma digital oficial VOC –,<sup>2</sup> e tem por objetivo último verificar a relevância do mesmo, fundamentalmente sob três perspetivas inter-relacionadas: sociológica, linguística e política. O trabalho pretende identificar casos concretos em que motivações políticas e sociolinguísticas influenciam a seleção e o uso de topónimos, nomeadamente a grafia destes; enquadrar os dados do VT com a tradição lexicográfica e com o prescrito desde 1911; entender de que forma está organizado o VT e qual a sua utilidade para o desenvolvimento de linhas toponímicas orientadoras que estejam em conformidade com as seguidas por organismos internacionais relativamente à toponímia, especificamente o UNGEGN;<sup>3</sup> finalmente, verificar quais as mais-valias decorrentes do VT que o justificam, e se efetivamente lhe conferem um posicionamento de base relativamente à padronização toponímica em língua portuguesa e ao desenvolvimento de aplicações de relevo em diferentes áreas do saber.

O VT é um vocabulário de especialidade que disponibiliza informação toponímica organizada de forma homogénea, cujos dados estão compilados de modo sistemático, após tratamento e desenvolvimento da informação disponível resultante dos trabalhos mais recentes levados a cabo pelos institutos nacionais de estatística de cada país participante.<sup>4</sup> Foram recolhidos dados dos referidos institutos por serem estes os

1 Segundo Castells (2005:20), “a sociedade em rede, em termos simples, é uma estrutura social baseada em redes operadas por tecnologias de comunicação e informação fundamentadas na microeletrónica e em redes digitais de computadores que geram, processam e distribuem informação a partir de conhecimento acumulado nos nós dessas redes”.

2 Vocabulário Ortográfico Comum da Língua Portuguesa, disponível em <http://voc.cplp.org/>, instrumento normalizador previsto no texto original do Acordo Ortográfico da Língua Portuguesa de 1990, Art. 2.º, alterado pelo Protocolo Modificativo ao Acordo Ortográfico da Língua Portuguesa: “Art. 2 – Os Estados signatários tomarão, através das instituições e órgãos competentes, as providências necessárias com vista à elaboração de um vocabulário ortográfico comum da língua portuguesa, tão completo quanto desejável e tão normalizador quanto possível, no que se refere às terminologias científicas e técnicas.”

3 Sobre o UNGEGN – United Nations Group of Experts on Geographical Names, ver ponto 4.7. desta dissertação.

4 Dados dos censos brasileiros de 2010 em <http://censo2010.ibge.gov.br/>; dados cabo-verdianos de 2010 disponíveis em <http://www.ine.cv/censo/censo2010.aspx>; dados moçambicanos de 2007 em [http://www.ine.gov.mz/operacoes-estatisticas/censos/censo-2007\\_](http://www.ine.gov.mz/operacoes-estatisticas/censos/censo-2007_); dados portugueses de 2011 disponíveis em [http://censos.ine.pt/xportal/xmain?xpid=CENSOS&xpgid=censos2011\\_apresentacao\\_](http://censos.ine.pt/xportal/xmain?xpid=CENSOS&xpgid=censos2011_apresentacao_); dados são-tomenses de 2012 em <http://www.ine.st/2012.html>;

repositórios toponímicos mais abrangentes e bem sistematizados ao nível de cada país, já que na sua génese são utilizados, oficialmente, para efeitos estatísticos. Passíveis de aprimoramento, os dados obtidos em primeira instância foram cruzados com outras fontes de referência, convencionais, como dicionários e vocabulários, bem como com páginas de Internet oficiais de unidades com relevância administrativa.

Além da atestação dos topónimos e da referência às relações funcionais entre estes e os gentílicos deles derivados, no VT é feita marcação explícita das propriedades formais de cada topónimo, necessária para as funções prescritiva e normalizadora que o recurso pretende assumir. As relações hierárquicas entre topónimos, com ligação geralmente biunívoca entre um topónimo e um gentílico deste dependente, são incluídas quando atestadas na tradição.<sup>5</sup>

Foram então desenvolvidas tabelas únicas com dados relativos aos referentes geográficos maiores de cada um dos países da CPLP participantes e aos restantes países e capitais do mundo, posteriormente tratados, aprimorados e organizados para conversão em *MySQL*, sistema relacional de bases de dados em que está organizado e armazenado o VT. Na base de dados, a informação relativa a cada um dos conjuntos encabeçados pelo nível «país» (para a CPLP) ou «mundo», para os países que não fazem parte da CPLP, outrora não relacionais, passam a comuns e hierarquicamente interligados.

Não está incluída no vocabulário informação etimológica, por não ser esse o escopo atribuído ao recurso, nem por ele almejado.

No âmbito da dissertação, foram identificados e analisados de forma aprofundada, quantitativa e qualitativamente, todos os conjuntos e subconjuntos de dados do VT, assim como as relações, hierárquicas ou derivacionais, existentes entre unidades toponímicas.

O Capítulo 1 contém a introdução, a metodologia adotada e a estrutura da dissertação.

O Capítulo 2 introduz e desenvolve o conceito de «toponímia», sendo explorada a sua vertente sociolinguística. É também explicada a diferença entre endónimos e exónimos e algumas implicações daí decorrentes. Aborda-se o emprego de artigo associado a topónimos e a formação previsível de gentílicos.

No Capítulo 3 desenvolve-se a noção de normalização linguística e a sua relação com a toponímia. É apresentada a plataforma VOC, enquanto instrumento político, sociológico e linguístico.

O Capítulo 4 introduz o Vocabulário Toponímico – recurso do VOC. São abordadas as mais-valias gerais da padronização de referentes toponímicos, com referência ao UNGEGN, grupo de trabalho da ONU.

O Capítulo 5 apresenta alguns dados do Vocabulário Toponímico. Primeiro, os relativos aos países e capitais, com apresentação de casos que merecem particular atenção; depois, aqueles referentes às divisões administrativas dos países da CPLP já tratados.

No Capítulo 6 são dadas conclusões, assim como hipóteses de desenvolvimento de trabalho na área.

---

dados leste-timorenses de 2015 disponíveis em <http://www.statistics.gov.tl/pt/>.

5 Gentílicos expectáveis mas sem tradição lexicográfica não se encontram atestados no VOC; logo, não estão associados, nesta versão, aos topónimos que constituem a nomenclatura do VT, justamente pela inexistência de fontes de referência que os validem. Assume-se que a ausência de atestação nestas fontes terá por base, entre outras motivações, a preferência por uma expressão analítica comum “de TOP”. Os gentílicos previsíveis derivados de topónimos, enquanto concorrentes não preferenciais de expressões analíticas, não são agramaticais por definição.

## 2. TOPONÍMIA

A toponímia é um ramo da onomástica, disciplina que estuda os nomes próprios, dividindo-os em dois grandes grupos, o dos antropónimos (nomes de pessoas) e o dos topónimos (nomes de lugares). A palavra «toponímia» deriva de «topónimo», que se formou a partir dos radicais gregos *τόπος*, transliterável para «topos», e *ὄνομα*, que tem «onoma» por forma transliterada. Os radicais «topos» e «onoma» significam, respetivamente, «lugar» e «nome», pelo que o equivalente sintagmático da palavra «topónimo» será «nome de lugar».

«Toponímia» remeterá, *grosso modo*, para o «estudo dos nomes de lugares», podendo dizer respeito a um conjunto complexo de lexemas bem delimitado – como é o caso do VT –, composto por subconjuntos relacionais; a uma simples lista de topónimos associados a uma restrita realidade política, lista que será sempre, ela própria, subconjunto de um grupo de topónimos maior que poderá ou não estar convenientemente estudado (i.e., “a toponímia de um país”; “a toponímia de um distrito”; etc.); ou a um repositório de dados maioritariamente etimológicos associados a um subconjunto específico de topónimos.

As definições de «topónimo» e «toponímia» surgem nos dicionários de forma relativamente consensual, ainda que existam obras com aceções distintas das que aqui se apresentam. Uma interpretação própria que tenha por base estas fontes será, crê-se, transversalmente aceite. A informação disponibilizada em quatro dicionários de referência, dois do Brasil (Ferreira, 1999; e Houaiss, 2001) e dois de Portugal (Casteleiro, 2001; e Costa 2004), uniformizada estruturalmente para conveniência na leitura, torna mais clara a prática comum adotada em dicionarística de língua portuguesa relativamente à definição dos termos «topónimo» e «toponímia»:<sup>6</sup>

**topônimo:** *s.m.* nome próprio de lugar. Ex. Europa, Espanha, Amazonas, Pará, Brasília, Maceió, Serra do Mar, Solimões. De *top(o)-* + *ônimo*.

**toponímia:** *s. f.* Estudo lingüístico ou histórico da origem dos topônimos. De *topônimo* + *ia*. (Ferreira 1999:1689)

**topônimo:** *adj.s.m.* 1 nome geográfico próprio de região, cidade, vila, povoação, lugar, rio, logradouro público etc. 2 origem de um nome geográfico. De *top(o)-* + *-ônimo*.

**toponímia:** *s.f.* 1 parte da onomástica que estuda os nomes próprios de lugares. 2 lista, relação de topônimos. 3 estudo etimológico e/ou histórico sobre os topônimos [Compreende diversas subdivisões, como corônimos, limnônimos, eremônimos, potamônimos, talassônimos etc.] 4 (*por metonímia*) livro (ou outro suporte) que contém tal relação ou estudo. ≈ toponimia. De *top(o)-* + *-onímia* ou *-onimia*, formada por analogia com palavra de origem grega; fonte histórica 1899 *toponýmia*. (Houaiss 2001:2735)

**topónimo:** *s.m.* nome próprio de um lugar. De *topo-* + *gr. ovoμα* 'nome para designar uma pessoa ou uma coisa'.

<sup>6</sup> Note-se que a tradição dicionarística para a língua portuguesa é ainda praticamente exclusiva de Portugal e Brasil.

**toponímia:** *s.f.* 1. sistema de nomes próprio de lugares; conjunto de topónimos ≈ *toponomástica*.  
2. estudo linguístico, histórico de nomes próprios de lugares; estudo dos topónimos. ≈ *toponomástica*. De *topo-* + *gr: ovoμα* 'nome para designar uma pessoa ou uma coisa' + suf. *-ia*. (Casteleiro 2001:3586)

**topónimo:** *s.m.* nome de uma localidade, de um lugar, de um sítio. Do grego *tópos*, «lugar» + *ónyma*, «nome».

**toponímia:** *s.f.* 1 conjunto de topónimos; 2 estudo da etimologia e explicação dos topónimos. De *tópos*, «lugar» + *-ónyma*, por *ónoma*, «nome», + *-ia*. (Costa 2004:1508)

No âmbito deste trabalho, tendo por base as quatro referências agora citadas mas desenvolvendo o escopo da definição que se adota, assume-se que um topónimo é um nome próprio (simples ou composto) usado para identificar inequivocamente uma – e uma só – unidade geográfica. Designa, portanto, unidades geográficas<sup>7</sup> enquanto entidades únicas no mundo, ainda que estas possam ser homónimas de outras. Deve desde já ser referida a existência de um vastíssimo número de topónimos possíveis nomeando um único lugar, decorrente do plurilinguismo que exorna e aformoseia o nosso planeta.

Quando falamos de toponímia, não podemos deixar de ter em linha de conta a frequente existência de artigos<sup>8</sup> coocorrentes com um topónimo, efetivando a hipótese de que estes são parte das suas propriedades formais, assim como é útil a referência a dados que não são parte dos topónimos, *per se*, embora com eles estejam relacionados: entre estes, destaca-se a previsível associação de informação gentílica aos topónimos, ou seja, o estabelecimento de uma relação quase sempre derivacional entre gentílicos<sup>9</sup> (igualmente designados como «etnónimos» ou «nomes pátrios») e topónimos.

## 2.1. Toponímia: perspetiva sociolinguística

A motivação que serve de fio condutor à nomeação de um lugar de determinada forma em detrimento de outra(s) tem, necessariamente, forte cunho cultural, podendo ser estudada ao nível da sociologia, inclusive nos casos em que é o resultado de decisões fundamentalmente políticas.

Entre as diferentes abordagens possíveis quando se trata de toponímia, a mais frequente remete, ainda assim, para a análise histórica, diacrónica, etimológica: para o estudo da origem e da evolução dos nomes de lugares. Outra abordagem, distinta mas que também tem presente a relevância dos dados históricos, tem por foco a criação de acervos lexicais para constituição de vocabulários de especialidade que são parte do léxico geral de uma língua, com especificidades maioritariamente sincrónicas. A elaboração de um vocabulário deste tipo deve seguir critérios rigorosos, com a expectável seleção de fontes fidedignas mas, também, com uma interpretação correta da realidade social do tempo presente, que tanto resulta da tradição como a

---

7 A relevância da dimensão das unidades toponímicas é desenvolvida no ponto 4.3.1. desta dissertação.

8 Sobre artigos, veja-se o ponto 2.3.1.

9 Sobre gentílicos, desenvolve-se o texto no ponto 2.3.2. desta dissertação.



transforma para os tempos vindouros, moldando-a conforme as necessidades de comunicação se vão atualizando.

Uma língua é uma realidade com duas faces indissociáveis: uma realidade mental e uma realidade social. Esta face social da língua dá origem ao estabelecimento de convenções a respeito dela, à determinação de usos linguísticos socialmente aceites, considerados mais corretos do que outros, isto é, dá origem à norma linguística, entendida neste contexto, de acordo com Mateus e Cardeira (2007), como a modalidade linguística que é escolhida por uma sociedade enquanto modelo de comunicação. (Correia e Ferreira 2013:297)

Os “usos linguísticos socialmente aceites” numa língua referidos em Correia e Ferreira (2013) – ou, se preferirmos, a realidade social associada a uma língua – incluem, claro, a toponímia. Existem designações alternativas (concorrentes ou complementares) para um número inaferrível de topónimos, cuja gramaticalidade advém do uso generalizado dos falantes, normalmente respaldado pelos dados que constam de fontes de referência. Assim, por mais que uma designação toponímica atribuída a um referente geográfico único seja preferida relativamente a outras possíveis e até, eventualmente, mais bem fundamentadas, o conceito de «tradição», que aporta um nível específico de normatividade, é parte nuclear de qualquer topónimo em uso, mesmo quando serve de justificação para um corte com a designação que se entende substituir. Ou seja, qualquer topónimo em uso tem associada uma noção de aceitabilidade construída ao longo do tempo, com maior ou menor força de tradição, que num extremo pouco marcado o identificará como sendo somente «possível» e no outro o terá como «exclusivo». Neste trabalho, assume-se a existência de uma forte relação entre grau de normatividade e peso da tradição, que ainda assim, pontualmente, pode não se verificar.

Aquando da nomeação de um lugar numa língua, num extremo admite-se a existência de diferentes topónimos para o designar – possuindo um deles nível de aceitação superior ou não – e no outro exclui-se todo e qualquer topónimo concorrente àquele que se tem como autorizado pela norma. Entre os extremos propostos, é possível atribuir diferentes graus, tais como: «adequado», que admitirá (ou não) a existência de outros topónimos para um mesmo lugar, sem que exista especificação quanto à sobreposição do topónimo identificado, em termos de aceitação, relativamente aos restantes; ou «preferível», nível em que também se admitirá diferentes designações, mas marcadamente não preferenciais.

É hoje relativamente simples compreender a relevância do estrito compromisso de manutenção e consolidação, por parte de uma língua pluricêntrica, de topónimos grafados tendo por base uma política internacional de corpo comum, balizada por princípios científicos e por instituições que permitam o seu desenvolvimento. Este papel, para a língua portuguesa, é assumido pela CPLP, especificamente pelo seu Instituto Internacional de Língua Portuguesa (IILP). Os estados-membros da CPLP, defendendo o património linguístico que partilham, reafirmam todo o interesse em possuir e divulgar topónimos em língua portuguesa que designem todos os lugares reconhecidos como seus pelas populações. Tal interesse não inviabiliza a

existência de topónimos grafados noutras línguas, necessariamente de contacto, mas reforça a apropriação de um nome de lugar numa língua que é oficial e que pode ser tida em conta como ferramenta disponível e desejável na identificação toponímica pelos falantes de português.

### 2.1.1. Fatores identitários e de representatividade social

Os fatores identitários que compõem a toponímia têm sempre uma base histórica bem vincada, quase sempre indissociável das noções de «naturalidade», de «pátria» ou de «soberania»: “a identidade nacional resulta antes de mais da perceção que os próprios cidadãos têm de formarem uma coletividade humana (...), dado que essa coletividade tem uma existência histórica” (Mattoso, 1998:3). A noção de identidade nacional, por exemplo, está presente no consciente de qualquer cidadão, mas este sentimento de pertença a uma coletividade é também partilhado ao nível local, por comunidades cuja dimensão, menor, é inversamente proporcional ao grau de proximidade entre pessoas.

A proximidade entre um indivíduo, um lugar e o nome do próprio lugar, sob uma perspetiva identitária, enquadra-se nas identidades primárias de Castells, “fundamentais para a construção da identidade coletiva” (Castells, 1999:28). São elas as identidades religiosa, nacional, étnica e territorial, interessando-nos particularmente esta última. Considerar um território como constituinte primeiro de um todo identitário complexo é revelador da importância da designação toponímica na construção do eu, mais não seja pelas propriedades particularizadoras que carrega. Um nome específico atribuído a um lugar resulta de uma articulação crítica relativamente a três fatores: “o homem que o produz; dentro de determinada cultura (de acordo com sua cosmovisão); situado em determinado espaço e em certa temporalidade.” Carvalhinhos (2008:12) sublinha a importância destes fatores acrescentando que “somente enquadrando o topónimo neste tripé é possível compreendê-lo e interpretá-lo como realmente é, e não apenas como componente de uma lista de nomes seguida do provável significado”.

### 2.1.2. Paisagem linguística e coexistência de línguas locais

À “visibilidade e saliência” de uma ou mais línguas em sinais públicos ou comerciais<sup>10</sup> existentes num território ou região, Landry (1997) chama «paisagem linguística». A marcação efetiva de fronteiras linguísticas é útil para a delimitação de território, mas serve, simultaneamente, de entrave às relações entre grupos linguísticos distintos, amplificando o sentimento de pertença associado a cada um deles e criando um distanciamento relativamente a quaisquer grupos tidos como «exteriores», em particular os geograficamente contíguos. Esta constatação torna claro que a paisagem linguística não possui uma função meramente informativa. A ligação entre toponímia e paisagem linguística é particularmente evidente em contextos multilingues.

No caso da língua portuguesa, que não raro convive com outras línguas locais nos espaços em que é oficial, a opção pela codificação de topónimos sob uma língua comum, no âmbito de uma política linguística

---

10 É feita a associação, nesta dissertação, entre o conceito «comercial» e a noção «de origem privada».

partilhada pelos países da CPLP, possibilita o desenvolvimento de e acesso a recursos únicos e homogêneos. As vantagens percebidas daí decorrentes incluem a identificação transnacional de topónimos e a associação direta da pertença de um lugar e das suas populações a uma comunidade maior. Um sistema toponímico sob o escopo de uma língua pluricêntrica não invalida, sublinhe-se, o respeito pelas línguas com as quais partilhe geografias, nem exclui sistemas alternativos de representação, equidistantes ou não.

### 2.1.3. Topónimos híbridos

Até à elaboração e disponibilização do VT, na maioria dos países de língua portuguesa não existia uma lista toponímica oficial. As existentes misturavam, frequentemente, topónimos escritos em várias línguas, seguindo preceitos heterogêneos quanto à fixação das formas e desatualizados quanto ao formato e padrões seguidos. Além da existência de topónimos diferentes para nomear um único lugar, outros viam-se atestados pelo recurso à amálgama de constituintes provenientes de duas ou mais línguas, tal como acontece com inúmeras formas do léxico geral não totalmente adaptadas à grafia do português. Estes casos podem ser designados como «topónimos híbridos». Dada a dimensão da língua portuguesa, sendo atualmente a mais falada no hemisfério sul e a quarta mais falada do mundo, de acordo com o Observatório da Língua Portuguesa,<sup>11</sup> a toponímia grafada em língua portuguesa terá inúmeras vantagens em respeitar uma norma ortográfica comum, bem como linhas orientadoras que procurem suprimir quaisquer fenómenos de variação puramente ortográfica.

O VT presta-se à normalização toponímica, através da disponibilização dos topónimos que atesta, grafados em língua portuguesa, procurando necessariamente resolver casos de potencial hibridismo, pelo que otimamente estará a atestação de topónimos híbridos restringida a zero ocorrências. Exemplos haverá, como «Porto Moresby», capital da Papua Nova Guiné, em que identificamos, à superfície, um topónimo híbrido. Na realidade, a forma atualmente usada em português decorre tão só de uma tradução do primeiro elemento, de «Port» para «Porto», e da manutenção do segundo elemento, por se tratar de um constituinte de origem antroponímica (de John Moresby). Assim, embora o topónimo seja híbrido no sentido estrito do termo (por possuir constituintes de línguas diferentes), um dos constituintes, sendo inicialmente um antropónimo, possui restrição de adaptação às regras ortográficas da língua de chegada, seguindo a prática comum para a língua portuguesa, refletida no AO90 (Base I, ponto 3): “[...] mantém-se nos vocábulos derivados eruditamente de nomes próprios estrangeiros quaisquer combinações gráficas ou sinais diacríticos não peculiares à nossa escrita que figurem nesses nomes: [...] *garrettiano*, de *Garrett*; *jeffersónia/jeffersônia*, de *Jefferson* [...]”

Como exemplo de um processo de adaptação pleno, observemos com tento o topónimo «São Marinho». Este topónimo passou por um período de volubilidade, com alternância de atestação entre a forma totalmente adaptada e uma outra forma, eventualmente ainda de uso bastante frequente entre os falantes de língua portuguesa, «São Marino». Pela possibilidade de adaptação sem constrangimentos maiores, tendo por base a

---

11 A língua portuguesa tem aproximadamente 261 milhões de falantes, superada apenas pelo mandarim (848M), pelo espanhol (399M) e pelo inglês (335M), e posicionada imediatamente acima do hindi (260M) e do bengali (189M). Dados de 15 de fevereiro de 2016, consultados em <http://observalinguaportuguesa.org/as-20-linguas-mais-faladas-no-mundo/> em abril de 2016.

já existente atestação do topónimo plenamente adaptado em fontes de referência e a desejabilidade de cassar topónimos híbridos, o VT atesta «São Marinho», com ambos os constituintes homónimos de palavras do léxico geral.

## **2.2. Outras questões com implicações sociológicas**

### **2.2.1. Endónimos e exónimos**

Um lugar pode ser nomeado em diferentes línguas, pelo que existem potencialmente inúmeras formas legítimas de designar o mesmo referente geográfico. Os topónimos podem ser classificados quanto à sua relação político-linguística com a língua em que são usados: são «endónimos» os topónimos referentes a lugares em que a língua oficial ou dominante é a que está a ser usada; são, pelo contrário, classificados como «exónimos» os topónimos referentes a lugares em que a língua que está a ser usada não tem estatuto de oficialidade. Assim, em português, são exónimos todos os nomes de lugares fora dos espaços de língua oficial portuguesa. Atente-se na definição de UNGEGN (1972) para «exónimo»: “[Um exónimo é] um nome geográfico usado numa determinada língua para uma entidade geográfica localizada fora da área onde essa língua tem estatuto oficial e diferindo na sua forma do nome usado na língua ou línguas oficiais da área onde a entidade geográfica está localizada.” UNGEGN (1972:5) [tradução própria]

O conceito aqui adotado, usado no âmbito do VT, determina que sejam tomados por endónimos todos os topónimos de países de língua portuguesa em português e como exónimos todos os outros. Nas palavras de Kerfoot (2006:129) um endónimo é sempre “um nome de lugar cujas autoridades nacionais estão autorizadas a padronizar”. Esta função homogeneizadora pode – deve acrescentar-se – transpor fronteiras, de forma consensual, sendo para tal necessário existir uma norma ortográfica comum em vigor nos países que decidam partilhar a referida estratégia, existência que se verifica para a língua portuguesa.<sup>12</sup>

De uma perspetiva sociológica, todo o trabalho realizado no sentido de atestar endónimos únicos em língua portuguesa no espaço da lusofonia tem importância acrescida, na medida em que reforça a relação entre o idioma e todos os seus utilizadores, e potencia uma aproximação, em termos culturais e identitários, entre as populações que a utilizam. Caso se enquadre a atestação de endónimos únicos em língua portuguesa numa perspetiva meramente economicista, torna-se também evidente que a existência de um repositório com endónimos comuns a toda a CPLP traz vantagens dificilmente alcançadas caso se mantivesse a elevada variação existente até aos dias de hoje, decorrente de medidas políticas tomadas unilateralmente por cada país para a sua atestação.

Se tivermos em conta a forma como os exónimos são incorporados numa língua, podemos enquadrá-los em, pelo menos, cinco subgrupos, seguindo Ormeling (1993):

---

<sup>12</sup> Importa, a este respeito, chamar a atenção para o importante conceito de «forma não adaptada», tratado no ponto 4, que é distinto deste.

- (i) Exónimos que resultam de uma transcrição<sup>13</sup>  
(e.g. *Македонија* → *Makedonija* → Macedónia)
- (ii) Exónimos fruto de uma adaptação  
(e.g. *Kobenhavn* → Copenhaga)
- (iii) Exónimos com desenvolvimento separado partindo de uma origem comum  
(e.g. *Colonia* (Köln) → 1. *Cologne* (francês) 2. *Keulen* (neerlandês) 3. Colónia (português))
- (iv) Exónimos que não advêm do endónimo  
(e.g. *Hellas* → Grécia)
- (v) Exónimos resultantes de uma tradução  
(e.g. *Saint John's* → São José – capital de Antígua e Barbuda)

Além destes, será pertinente o reconhecimento da existência de exónimos criados a partir da forma fonética, por exemplo nos casos de línguas cuja grafia é ou era desconhecida ou inexistente aquando da primeira codificação (e.g. Zaire).

Um exemplo como o da forma «Paris» corresponderá a *n* topónimos homógrafos que identificam o mesmo espaço geográfico mas que podem ser considerados ora endónimo(s) (e.g. [pa'Ri], em francês), ora exónimos (e.g. [p'Éris], em inglês; ou [pa'Riz], em alemão). Como se vê pela tipologia proposta por Omerling (1993), acima, é comum e possível a adaptação de exónimos a todas as línguas do mundo, em conformidade com cada sistema gráfico existente, não esquecendo o sistema fonológico de cada língua aquando da adaptação. Daí resultam diferentes topónimos, pelo menos um por língua, se considerarmos que cada um deles está associado a um sistema político-linguístico, mesmo que eventualmente haja formas homógrafas em diferentes línguas (e.g. Paris). Em casos como o de «Москва», em português «Moscovo», a transliteração ou adaptação é mesmo uma necessidade prática. O mesmo ocorre com exemplos mais próximos, como «Espanña», necessariamente adaptado para «Espanha». Assume-se pois como desejável, sempre que possível, a adaptação das formas dos exónimos às regras de escrita de uma língua, para uma adequada integração em recursos como, para a língua portuguesa, o VT.

### 2.2.2. Implicações político-diplomáticas da toponímia

Dadas estas limitações, a coexistência de endónimos e exónimos levanta problemas práticos e de ordem político-diplomática e sociolinguística em instituições internacionais, em que as línguas de trabalho, partilhadas por muitos países, não são apenas representativas dos países em que são oficiais, mas antes de todos os representados nessas instituições. Assim, em instituições internacionais a seleção de um endónimo

---

<sup>13</sup> Na verdade, estes são tipicamente casos de transliteração (neste caso em particular, há ainda um processo de adaptação, dado o valor de <c>).

na língua oficial usada no espaço onde se encontra o topónimo nomeado é preferencial. No entanto, está tradicionalizado para muitos topónimos o uso de um «nome convencional».<sup>14</sup>

Contrariamente ao que seria expectável para topónimos maiores, na escolha de nomes oficiais de cidades ou países pode não existir consenso quanto a uma forma toponímica exclusiva ou preferencial. A falta de consenso é normalmente de âmbito nacional, estando relacionada com o(s) endónimo(s) – e gentílicos deste(s) derivados – a adotar, mas, em alguns casos, pode ser alargada aos exónimos aceites pelo país que contém o lugar nomeado. Não é inexistente, portanto, a tentativa de imposição, por parte de um país, de exónimos homógrafos de um seu endónimo. Por regra, esta posição – mais protecionista, com motivações políticas, culturais e/ou ideológicas<sup>15</sup> – pretende ser reflexo da tradição assumida como nacional e serve como chancela de soberania, seja ela recém-adquirida ou outrora mais visível, agora entendida como obscurantizada e necessitando de reforço também pela via toponímica.

O caso de Timor-Leste é exemplo desta prática. As autoridades leste-timorenses defendem um topónimo único para nomear a sua nação, sublinhando que «Timor-Leste» deverá ser topónimo exclusivo, independentemente da língua em que seja empregue. O mesmo defendem os responsáveis da Costa do Marfim, que têm a ambição de ver o seu país conhecido universalmente por «Côte d'Ivoire», e alguns responsáveis indianos, que reclamam a substituição universal de «Bombaim», percebido como resquício do colonialismo europeu, por «Mumbai», retroconstruído a partir do nome de uma deusa.

Noutros casos, é o retorno a uma forma antiga que motiva a alteração formal, por decreto, de topónimos há muito em uso por outros que os terão antecedido. Tal prática foi comum aquando da independência das antigas colónias portuguesas, nomeadamente Moçambique, em topónimos como «Pemba», que durante parte do período colonial foi conhecida por «Porto Amélia».

Importa igualmente dar conta de um exemplo particular e muito recente relativo ao plebiscito popular para consagração por decreto da forma «acreano» como gentílico do topónimo Acre, estado do Brasil. Este gentílico corresponde, segundo os promotores da iniciativa, à forma mais empregue na oralidade, contrariando a disposição do VOLP, oficial no Brasil, que apenas atestava a forma regular «acriano».

---

14 Entende-se por «nome convencional» toda a designação toponímica cujo principal propósito seja o uso internacional não restrito a uma língua, tradicionalmente partilhado por diferentes línguas e usado em instituições com incumbências a nível continental ou global (e.g. «Grécia», «Greece», «Grèce» e não «Hellas»).

15 Interessa também dar conta de alguns casos de nomes de lugares, nomeadamente de ruas, em que um estado, como forma de represália ou de pressão política relativamente a outro, utiliza uma designação associada a algo ou alguém que é mal-amado pela nação visada. Por exemplo, “em 1984, o segmento da 16.ª Rua, Washington DC, onde a embaixada soviética estava localizada, foi renomeado «Andrei Sakharov Plaza», em homenagem ao físico nuclear vencedor de um prémio Nobel, preso soviético dissidente”: <https://www.theguardian.com/cities/2016/jun/28/streets-new-names-airbrush-politics-renaming-roads> (artigo consultado em maio de 2016).

## 2.3. Outros dados linguísticos

### 2.3.1. O uso de artigo em topónimos

No campo da toponímia, uma das dúvidas mais frequentes entre os falantes de língua portuguesa prende-se com a validade do emprego de artigo em topónimos específicos, normalmente dúvida nos casos em que existe pouca familiaridade entre o topónimo em questão e o utilizador do idioma. O VT, ao incluir a marcação explícita relativa ao uso de artigo para os topónimos que atesta, colmata a inexistência de recursos completos e homogêneos com este tipo de informação. Os países e capitais, assim como as unidades maiores de cada país da CPLP, dispõem já desta informação no VT, que previsivelmente será alargada às unidades menores. Os dados disponibilizados, além da utilidade para o público em geral, permitem tirar algumas conclusões relativamente à motivação subjacente ao emprego de artigo nos topónimos.

Um topónimo pode ou não ser obrigatoriamente antecedido de artigo. Enquanto uma grande parte dos antropónimos em língua portuguesa é facilmente associada aos géneros masculino ou feminino, mesmo sem artigo realizado gráfica ou foneticamente, no caso dos topónimos tal não se verifica: não é possível inferir o género de um topónimo sem o auxílio de um artigo realizado. No âmbito deste trabalho importa clarificar a peculiaridade do emprego de artigo nos topónimos.

Admita-se que o artigo constitui uma das propriedades intrínsecas não só dos antropónimos mas também dos topónimos. Com base nesta premissa, é admissível formular uma regra em que se assumam que todo e qualquer topónimo possui um artigo associado, embora em alguns casos esse artigo desempenhe mero papel funcional, não tendo realização gráfica/fonética. No entanto, a existir nos topónimos, o papel funcional dos artigos teria de ser suficiente para a atribuição de género em todos os casos, o que não se verifica: do ponto de vista do uso, qualquer topónimo que não aceite a realização gráfica/fonética de artigo está privado de qualquer marca de género, sendo neutro por definição.

Compreende-se neste trabalho que o artigo definido fará tendencialmente parte das propriedades formais dos topónimos, desempenhando simultaneamente uma função de reforço na identificação inequívoca de um referente único, nos casos em que haja homofonia, ainda que não seja esta a sua principal funcionalidade. Os topónimos homónimos de nomes comuns não selecionam, necessariamente, artigo, embora seja assumido, de acordo com a tradição, que o uso obrigatório ou preferencial de artigo em topónimos e antropónimos está diretamente relacionado com o facto de o nome próprio ser derivado de um nome comum, conforme é asseverado em Cunha (1984:225):

“[...] hoje, há mesmo um grande número de nomes próprios que exigem obrigatoriamente o acompanhamento do artigo definido. Entre essas razões, [...] a de ser o nome próprio originariamente um substantivo comum, construído com o artigo: a Guarda; o Porto; o Cairo (árabe *El-Kahira* = a vitoriosa); o *Havre* (francês *Le Havre* = o porto).”

Analisando os dados do VT, chega-se à conclusão de que a generalização tida como válida na tradição poderá não estar correta, havendo inúmeros casos que contrariam a regra apresentada em Cunha (1984). Nesta dissertação apenas se dá conta da fragilidade da generalização. Assume-se, por agora, que um topónimo grafado em língua portuguesa terá um género associado quando é obrigatório o uso de artigo na sua nomeação (e.g. a Etiópia [-MASC]), ou quando existe uma preferência pelo seu uso (e.g. em/na França [-MASC]), e que é essa a motivação e a função principal do artigo (independentemente de o topónimo ter ou não origem num nome comum seu homógrafo): a atribuição de género. Desta forma, torna-se mais simples identificar os topónimos que terão género – ou seja, aqueles em que um artigo é, ou pode ser, usado aquando da sua referência –, pela necessária concordância flexional.<sup>16</sup> Atentemos nos exemplos seguintes:

a. **As** Filipinas [-MASC,+PLU] **são** magníficas.

b. **Os** Açores [+MASC,+PLU] **são** magníficos.

Já os topónimos que não têm marca de género explícita, assumidos normalmente como sendo neutros, não terão concordância para verificar, logo, não pode, à superfície, ser-lhes atribuído género (c., d.). Se o topónimo em questão for parafraseado por «o país» ou «a cidade», por exemplo, é naturalmente empregue o género do nome comum que o substitui (c., d.). A flexão concorda também com o nome comum quando este é omitido, não devendo ser associada ao topónimo pelo simples facto de o artigo lhe ser contíguo (e.) ou de existir um adjetivo que aparenta flexionar com o topónimo (f.):

c. Portugal [Ø,-PLU] é um ótimo destino turístico.

c.' **O** país [+MASC,-PLU] tem vindo a apostar nesta área estratégica.

d. Dili [Ø,-PLU] é a capital de Timor-Leste.

d.' **A** cidade [-MASC,-PLU] tem muita margem de progressão.

e. **A** [cidade de] Díli [Ø,-PLU] dos anos 90 era bem diferente.

f. [A cidade de] Maputo [Ø,-PLU] é linda.

Nos casos em que se opte por omitir um artigo que é gramatical, a flexão, tanto em género como em número, mantém-se (g.-g.). A exceção dirá respeito a títulos jornalísticos, publicidade ou questões de estilo, em que se admite a omissão de um artigo «obrigatório», potenciando variação quanto ao número (h.-h.). Caso o artigo seja grafado, a concordância é obrigatória, sem exceção (i.-i.).

g. [O [estado do]] Amazonas, agora mais desenvolvido, **elegeu** um novo governador.

g.' \*[O [estado do]] Amazonas, agora mais desenvolvidos, **elegeram** um novo governador.

h. [**As**] Maldivas **são** destino de eleição.

---

<sup>16</sup> Há, ainda assim, casos em que um artigo apenas aparenta concordar em género com um topónimo, podendo a constatação mais óbvia aqui descrita não ser aplicável (exemplos e.-f.).



- h.' [As] Maldivas é destino de eleição.
- i. As Maldivas são destino de eleição.
- i.' \*As Maldivas é destino de eleição.

Quanto à motivação para a ausência de artigo em alguns topónimos, Garnelo (2001) *apud* García Arias (1995:81) chama a atenção para uma ocorrência relevante: “Outra característica da estrutura dos nomes que pode ser importante é a presença ou ausência de artigo; diz García Arias (1995:81) que ‘pode trabalhar-se, como hipótese aproximativa, com a ideia de que os topónimos formados de antigos antropónimos, se não usam artigo, podem ter uma grande antiguidade’”. Esta ideia reforça, no mínimo, a generalizada opção pela ausência de artigo verificada na nomeação de hagiónimos ou hierónimos (nomes de santos ou nomes sagrados, respetivamente). Daqui se torna mais plausível que os topónimos derivados de nomes de santos, de deuses ou de entidades sagradas terão o registo antigo como justificação para a ausência de artigo na sua nomeação, quando o uso nominal enquanto antropónimos não requeria artigo para aferir o seu género.<sup>17</sup> Mas além da questão da datação, outras motivações – como a veneração ou simples reverência – contribuirão para a generalizada ausência de emprego de artigo nestes casos. No caso de outros topónimos formados de antigos antropónimos, que não hagiónimos ou hierónimos, a ausência de um padrão suficientemente abrangente associado a datações exatas de primeiros registos não permite aferir a relação entre ausência de artigo e antiguidade, razão pela qual se poderá assumir que a formulação em Garnelo (2001) *apud* García Arias (1995:81) tenderá a ser produtivamente substituída por: “a grande antiguidade dos hagiónimos e hierónimos, formados de antigos antropónimos, autoriza a ausência de artigo”.

Ainda de acrescentar a previsível ocorrência de artigos com nomes de países. Dos 194 países reconhecidos pelas Nações Unidas, somente em 33 não é permitido o uso de artigo. Em 31 outros topónimos existe alternância relativamente ao seu uso (a nível nacional ou entre variedades nacionais). Segundo Bosque e Demonte (1999:4877): “Los nombres que denotan territorios presentan un comportamiento bastante variado. Los que acaban en *-a* son generalmente femeninos (*la Australia, la Grecia, la India y la Rusia del siglo XIX*), mientras que los demás son masculinos (*el Afganistán, el Chile, el Malawi, el Portugal y el Salvador del siglo XX*). Para o espanhol é, portanto, apresentada a teoria de que os nomes de territórios são, por regra, femininos quando terminam em *-a(s)* e masculinos com todas as outras terminações. Em língua portuguesa, os nomes de países femininos têm praticamente sempre uma terminação *-a(s)* (inclui a terminação do primeiro constituinte, no caso de topónimos compostos), exceto nos casos da Guiné, da Guiné Bissau e da Guiné Equatorial.<sup>18</sup> Se o ponto de partida for a terminação do topónimo, e não o género, como sugere Bosque e Demonte (1999), em 111 topónimos terminados em *-a(s)*, 90 têm género feminino e em 21 ora é simplesmente vedado o emprego de artigo, ora existe variação (a nível nacional ou entre variedades nacionais). Relativamente às restantes terminações, tendencialmente remetem para o género masculino,

---

17 Esta função era desempenhada pelos adjetivos «são», «santo», «santa», etc.

18 O nome «Guiné», enquanto topónimo ou constituinte de topónimo, se tivermos em conta os dados do Corpus do Português (<http://www.corpusdoportugues.org/>), passa a ter artigo feminino obrigatório apenas a partir do século XX, sendo usadas estruturas sintáticas sem marcação explícita de género até então (e.g. [...] em Guiné).

ainda que exista mais variação entre topónimos com género masculino e aqueles em que o género não pode ser aferido (e.g. Senegal [+MASC], mas Portugal [Ø]).

Tendo por base estes dados, é válido afirmar que o comportamento relativo à atribuição de género nos nomes de países em língua portuguesa se aproximará daquele assumido em Bosque e Demonte (1999) para o espanhol. Mas relativamente à língua portuguesa será igualmente pertinente a constatação de que os nomes de países que em todas as variedades da língua não admitem artigo constituem exceção (são 30 em 194: 15,46%). Olhando para topónimos como «Espanha», «França» e «Itália», e para a recente aceitabilidade de emprego de artigo aquando da sua nomeação, que não será desenvolvida neste trabalho, admite-se que a mesma decorrerá de uma regularização através de um processo de analogia sincrónica. Isto é, que a terminação *-a* destes topónimos frequentes (o mesmo acontece com o topónimo «Inglaterra», por exemplo) influencia a aceitabilidade e a daí decorrente gramaticalidade do emprego do artigo, o que reforça a já referida relevância da terminação para a definição de género.

### 2.3.2. Formação de gentílicos

Os gentílicos, em língua portuguesa, são formados normalmente por um processo de sufixação. Posto de forma simples, do encontro entre um topónimo e um sufixo nasce um gentílico. Os sufixos que permitem a ocorrência deste processo estão bem identificados pela tradição, possuindo, ainda assim, diferentes graus de produtividade. Por um lado existem dois sufixos mais produtivos e previsíveis, *-ense* e *-ês* – que mais não são do que duas distintas ocorrências atuais derivadas de um único sufixo do latim, *-ensis* –, e sufixos relativamente frequentes, ainda que menos do que *-ense* e *-ês*, como por exemplo *-ano*, *-eiro* e *-ão*; por outro lado temos aqueles sufixos a que Vasconcellos (1966:378) chamou “sufixos raros”:<sup>19</sup> *-eta*, *-ato*, *-oto*, *-engo*, *-ico*, *-eno*, *-ejo*, *-enho*, *-ista*, *-eno*, *-éu* são alguns exemplos.

Casos há em que um gentílico em uso não é derivado do topónimo atual a que está associado, mas de uma forma toponímica ancestral, com a qual mantém uma estreita relação. «Escalabitano» (Santarém), de *Scalābis*, nome antigo, é exemplificativo deste tipo de ocorrência. Vasconcellos (1966) dá conta dos casos com mais visibilidade à época dos seus escritos, ainda hoje, na sua maioria, válidos em Portugal:

Certos nomes derivam dos nomes antigos das povoações, ou até supostos primitivos, isto é, de latinizações ou helenizações de nomes modernos como: *Flaviense* (Chaves), de *Aquae Flaviae*, nome antigo; *Vimaranense* [(Guimarães)], de *Vimaranis*, nome medieval; *Coliponense* (Leiria), de *Collipo*, nome antigo; *Albicastrense* (Castelo Branco, alatinado em *Castrum Album*); *Calipolense* (Vila Viçosa, helenizado em *Καλλιπολις*, «bela cidade ou vila»). (Vasconcellos 1966:379)

Gentílicos de origem antiga mantêm-se por vezes exclusivos, mas em alguns casos passam a preferíveis ou adequados, quando uma nova forma gentílica derivada do topónimo em uso ganha aceitação e passa a conviver com a designação anterior. Iniciando-se este processo, a evolução tenderá, previsivelmente, a retirar progressivamente aceitabilidade ao gentílico antes tido como exclusivo, tornando-o arcaísmo. Estes

---

<sup>19</sup> Entendidos como raros quando associados à formação de um gentílico.

gentílicos, em uso mas excepcionais, devem ainda assim ser utilizados nas obras que referenciem os topónimos a que estão associados, ou que deles tratem. Em alguns casos, gentílicos paralelos coexistem durante longos períodos de tempo sem que um deles seja assumido como mais adequado.

Acontece que para uma mesma povoação há não raro derivados pluriformes: *Bracarense* e *Braguês*; *Bragançano*, *Braganção* (antigo), e *Brigantino*. Um é formado da palavra antiga (Bracara); outro da palavra moderna (Braga). [...] Lisboa tem: *Olisiponense* (Olisipo), *Lisbonense* (Lisbona), *Lisboeta* (Lisboa), e em certos textos *Lisboês*, *Lisbonês* e *Ulixbonense*. (Vasconcellos 1966:379)

Relativamente ao emprego de hífen nestes elementos toponímicos, a regra a seguir é bastante transparente: os gentílicos compostos<sup>20</sup> derivados de topónimos são sempre hifenizados, seguindo a tradição, espelhada no outrora vigente AO45 e consubstanciada, atualmente, pelo AO90, com o acréscimo de exemplos: “[...] mato-grossense, norte-americano,<sup>21</sup> porto-alegrense, sul-africano [...]”.<sup>22</sup>

[...] Entretanto, os derivados directos dos compostos onomásticos em referência, tanto dos que requerem como dos que dispensam o uso do hífen, exigem este sinal, à maneira do que sucede com os derivados directos de compostos similares do vocabulário comum. Quer dizer: do mesmo modo que se escreve, por exemplo, *bem-me-querzinho*, *grande-oficialato*, *grão-mestrado*, *guarda-moria*, *pára-quedista*, *santa-fêzal*, em harmonia com *bem-me-quer*, *grande-oficial*, *grão-mestre*, *guarda-mor*, *pára-quedas*, *santa-fé*, deve escrever-se: *belo-horizontino*, de *Belo Horizonte*; *castelo-vidense*, de *Castelo de Vide*; *espírito-santense*, de *Espírito Santo*; *juiz-forano*, de *Juiz de Fora*; *ponte-limense*, de *Ponte de Lima*; *porto-alegrense*, de *Porto Alegre*; *são-tomense*, de *São* [ou *S.*] *Tomé*; *vila-realense*, de *Vila Real*.<sup>23</sup>

Esta redação vem no seguimento do prescrito nas Conclusões Complementares do Acordo de 1931, segundo as quais o hífen teria de ser empregue nos derivados de compostos onomásticos formados por justaposição:

Limitação do emprego do hífen, de acordo com o uso tradicional e corrente, em compostos do vocabulário onomástico formados por justaposição de palavras (*Vila Real*, *Belo Horizonte*, *Santo Tirso*, *Rio de Janeiro*, porém *Montemor-o-Novo*, *Grã-Bretanha*, *Áustria-Hungria* [...]); e emprego do mesmo sinal nos derivados de compostos onomásticos desse tipo (*vila-realense*, *belo-horizontino*, *austro-húngaro*).<sup>24</sup>

---

20 Gentílicos compostos possuem dois ou mais elementos lexicais.

21 «Norte-americano», que é gentílico paralelo de «estado-unidense» – empregue no Brasil – para o topónimo «Estados Unidos da América», deriva dos constituintes do nome do continente que contém o país (América do Norte), e não dos constituintes do nome do país. Este gentílico é então associado a um país e a um continente, ainda que derive do topónimo que identifica o continente.

22 Acordo Ortográfico de 1990, Base XV, ponto 1.

23 Bases Analíticas do Acordo Ortográfico de 1945 – Documento n.º 2, 28.

24 Conclusões Complementares do Acordo de 1931 – Segunda Parte, 28.



### 3. NORMALIZAÇÃO LINGUÍSTICA E RECURSOS GERAIS

A normalização linguística, que consiste na sistematização e veiculação das regras que regem um sistema linguístico, engloba a ortografia de uma língua. Como ferramenta para esta normalização, são utilizados vários recursos lexicográficos, nomeadamente dicionários monolíngues e vocabulários gerais ou de especialidade. Outros instrumentos, como a imprensa, obras de ficção ou registos escritos que constituem uma paisagem linguística, terão como modelo, previsivelmente, os dicionários e vocabulários, tidos como fonte de referência.

Os dicionários desempenham o papel de normalizadores de ortografia pelo simples facto de escolherem uma determinada forma ortográfica para representar a entrada lexical, mas essa informação é apenas uma entre muitas outras que os dicionários fornecem. Ao contrário, os vocabulários ortográficos são os instrumentos por excelência da regulação da ortografia, uma vez que é essa a informação principal que contêm: os vocabulários ortográficos tradicionais são constituídos por uma mera lista alargada de entradas, grafadas de acordo com a ortografia em vigor, podendo ou não conter indicação da categoria morfossintática e, em alguns casos, de pronúncia ou de irregularidades de flexão. (Correia e Ferreira 2013:315)

#### 3.1. Norma ortográfica e variação lexical

A língua portuguesa, à semelhança de grande parte das línguas naturais, possui um sistema gráfico utilizado, *grosso modo*, para representar produções orais com significado. *Grosso modo* porque, quando suficientemente afastada do registo oral, a escrita pode não desempenhar um papel meramente representativo, sendo legítimo considerá-la um sistema paralelo, não necessariamente complementar à oralidade. Até, em última instância, uma língua diferente. Esta abordagem, embora válida, é propositadamente secundarizada neste trabalho, que abraça a máxima efetivação da transparência entre grafia e oralidade. É, também, dificilmente aplicável à língua portuguesa contemporânea (cada vez mais afastada de um modelo que promova a manutenção da opacidade entre os dois registos), pelo reforço, hoje, da matriz fonológica que serve de alicerce à ortografia do português.

São muitas as relevantes aplicações práticas da representação escrita, nomeadamente aquelas que decorrem do carácter não efémero deste registo. Tal como na oralidade importa que uma mensagem seja corretamente percecionada, compreendida e interpretada segundo a conceção original de quem a transmite, também na escrita se assume como desejável este processo de transmissão de informação. Se na oralidade, para uma mesma língua, existem não raro diferentes variedades, o mesmo se traduz na escrita. E tal como cada falante tem um perfil normativo relativamente àquilo que diz e ouve – nem sempre correspondente à(s) norma(s)-padrão –, também quem escreve ou lê aplica regras que considera serem a «norma culta», regras que, ainda assim, podem corresponder a uma norma própria do escrevente, com todos os desvios passíveis de serem contemplados.

A grande diferença entre os dois meios de transmitir informação (oralidade e escrita) reside no facto de a escrita possuir uma «norma ortográfica» oficializada. A norma ortográfica para a língua portuguesa é, como já se referiu, de base fonológica, i.e., tem em conta a variação própria do registo oral e tenta codificá-la da forma mais abrangente e menos complexa possível. Não esquece, também, traços etimológicos consagrados pela tradição. Está plasmada em texto legal, sendo qualquer desvio ao seu ditame facilmente reconhecido e reprovado, tendo por base a força de lei. Na oralidade ocorre um processo normativo semelhante mas para este registo não existe um «vocabulário oral» que consolide efetivamente uma «norma oral», tarefa dificultada pela variação existente. O mais próximo que existe na prescrição de uma «norma oral» decorre, primeiro, das transcrições fonéticas atestadas em Casteleiro (2001), com a natural estreiteza que uma única transcrição fonética por palavra “segundo a norma culta, aproximada, de Lisboa e Centro do país” representa; e, depois, daquelas atestadas em Costa (2004),<sup>25</sup> que possuem o mesmo tipo de limitação.

Tradicionalmente, qualquer norma ortográfica em vigor num país, decorrente de uma «reforma ortográfica» ou de um «acordo ortográfico», é lei, sendo fruto da discussão e redação feita por especialistas, ainda que com um forte cunho político, na medida em que é aplicada por um estado soberano, como instrumento de normalização da escrita num espaço delimitado politicamente. Uma norma ortográfica, com carácter oficial, é desejavelmente acompanhada de um vocabulário ortográfico que tem por função disponibilizar todas as palavras em uso numa língua, seguindo uma interpretação estrita das regras esplanadas no texto legal que se encontre em vigor. O vocabulário oficial decorrente das regras do AO90 foi naturalmente previsto, ainda que a sua publicação passados dois anos da assinatura do acordo, como foi inicialmente previsto, não se tenha concretizado, por constrangimentos de ordem política. Está somente agora, em 2016, disponível em rede na sua primeira versão.<sup>26</sup>

Em casos específicos de línguas pluricêntricas, isto é, de línguas que possuem várias normas-padrão (de que é exemplo o português), uma norma ortográfica pode ser partilhada por mais do que um país. Idealmente, todos os países que partilhem uma língua oficial terão uma – e apenas uma – norma ortográfica para essa língua, já que esta é a base fundamental para a sua representação escrita. Não devem, portanto, ser confundidas as noções de «norma ortográfica» e de «léxico». A primeira corresponde a um metassistema linguístico em que estão plasmadas as regras ortográficas que regem a escrita de uma língua; a segunda é equivalente ao conjunto de formas lexicais de uma língua, isto é, possíveis na própria língua e reguladas, no caso do registo escrito, pela norma ortográfica.

A variação lexical é comum não só entre diferentes países que partilham a mesma língua, como em cada um desses países, internamente. Esta variação é salutar se tivermos em consideração a riqueza linguística de uma língua e a abrangência cultural inerente à diversidade permitida pela existência de palavras diferentes empregues na designação de «objetos» iguais ou semelhantes. Já a existência de mais do que uma norma ortográfica para uma só língua cria variação puramente ortográfica. Entendendo-se a língua portuguesa como entidade una, independente de quaisquer fronteiras políticas – logo, tratável enquanto «unidade soberana» –,

---

25 Em 2010, a reedição desta obra incluiu também transcrição fonética.

26 Ver ponto 3.2. desta dissertação.

poder-se-á permitir uma analogia com a existência de uma Constituição da República Portuguesa (aqui a representar «a mesma língua») que possua redações e artigos diferentes consoante a cidade em que é aplicada (entenda-se, «diferentes normas ortográficas»). A ser assim, estaríamos perante «constituições da República» para uma única «unidade soberana» que não seriam, em circunstância alguma, a mesma. Pode argumentar-se que no caso da Constituição da República Portuguesa estamos perante um instrumento nacional, com todas as implicações de cariz patriótico – ou até nacionalista – que daí podem advir, e que a norma ortográfica deverá ser balizada da mesma forma, ao nível estritamente nacional: uma «ortografia nacional», por oposição a uma «ortografia geral» da língua. Mas a realidade é que a «norma ortográfica» deve servir de instrumento de base para a ortografia da língua, não para uma variedade ortográfica nacional, embora contemple as especificidades das variedades que possam existir. Este facto é tão evidente que, se assumirmos, a título de exemplo, o português de Angola e o português da Guiné-Bissau como variedades da mesma língua, será tudo menos lógico criar normas ortográficas díspares, defendidas tendo por base variação de âmbito lexical.

Reforçando esta ideia, com o objetivo de esclarecer dúvidas remanescentes, um recurso digital que disponha de um chamado «corretor ortográfico», de que é exemplo o *Word*, disponibilizado no *Microsoft Office 2016*, poderá dar aos utilizadores como opção duas ou mais variedades de uma mesma língua, como é o caso do português de Portugal e do português do Brasil, já que os dados lexicais reconhecidos por cada uma das opções selecionadas serão necessariamente diferentes. Mas a norma ortográfica subjacente a estes quer-se a mesma, almejando criar, numa análise empírica da base ortográfica, natural homogeneidade ao nível da representação metalinguística, e não o seu oposto.

### 3.1.1. Reformas e acordos ortográficos

O AO90 é o acordo ortográfico atualmente em vigor para a língua portuguesa nos países que o ratificaram – todos os países da CPLP, à exceção de Angola, Moçambique e Guiné Equatorial, por motivos políticos distintos. O AO90 surge no seguimento das reformas e acordos realizados desde 1911, ano da entrada em vigor do «Formulário Ortográfico», primeiro documento oficial para a ortografia do português. Embora o Brasil não adotasse este formulário, em 1915 a Academia Brasileira de Letras harmoniza a escrita da variedade brasileira com a portuguesa, decisão que viria a ser revogada em 1919.

Entre avanços e recuos, em 1931 é aprovado o primeiro acordo ortográfico entre o Brasil e Portugal, com o intuito de unificar e simplificar a ortografia da língua portuguesa. Este acordo acabaria por não ser posto em prática. Em 1943 é redigido o «Formulário Ortográfico de 1943», na primeira «Convenção Ortográfica» entre o Brasil e Portugal; e em 1945 um novo Acordo Ortográfico torna-se lei em Portugal, não sendo ratificado pelo governo brasileiro, que continua a aplicar as normas do formulário de 1943.

Na década de 70 do século XX são promulgadas alterações no Brasil (1971) e em Portugal (1973) que reduzem as divergências ortográficas entre os dois países. Em 1986 é apresentado, no Rio de Janeiro, o «Memorando Sobre o Acordo Ortográfico da Língua Portuguesa», resultado do encontro entre os então países de língua oficial portuguesa, ainda que o Acordo Ortográfico de 1986 nunca viesse a ser aprovado. Em

1990, a Academia das Ciências de Lisboa e a Academia Brasileira de Letras elaboram a base do «Acordo Ortográfico da Língua Portuguesa», modificada por dois protocolos (1998) e (2004), passando o AO90 a entrar em vigor pela ratificação de três dos estados-membros da CPLP. Em 2006, Brasil, Cabo Verde e São Tomé e Príncipe ratificam o AO90. Em 2015 entra em vigor o AO90 em Portugal e em Cabo Verde. Em janeiro de 2016, o Brasil aplica na sua plenitude o AO90.<sup>27</sup>

### 3.2. VOC: enquadramento e papel normalizador

Tendo o AO90 sido assinado por todos os países da CPLP, e depois da sua promulgação em Portugal por parte do então presidente da República Portuguesa, Aníbal Cavaco Silva, em julho de 2008 – seguindo a aprovação do Segundo Protocolo Modificativo ao Acordo Ortográfico da Língua Portuguesa na Assembleia da República Portuguesa, em maio desse mesmo ano –,<sup>28</sup> o *Plano de Ação de Brasília para a Promoção, a Difusão e a Projeção da Língua Portuguesa* viria a recomendar que:

“Os Estados Membros que ainda não possuem Vocabulário Ortográfico da Língua Portuguesa [devem desenvolver] 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”.

Iniciou-se, então, o desenvolvimento do referido vocabulário ortográfico comum para os países da CPLP, o Vocabulário Ortográfico Comum da Língua Portuguesa (VOC), instrumento legal previsto para a aplicação plena do AO90. O VOC é uma plataforma digital de distribuição de dados lexicais que desempenha um papel de normalização ortográfica para a língua portuguesa, agregando os vocabulários nacionais dos países da CPLP numa base de dados lexicais comum e de acesso totalmente livre. É reflexo das fontes, da frequência e das propriedades específicas das formas mais representativas de cada uma das nações participantes, incluindo, à data: o VOLP (Vocabulário Ortográfico da Língua Portuguesa – Brasil); o VOCALP (Vocabulário Cabo-Verdiano da Língua Portuguesa); o VOMOLP (Vocabulário Moçambicano da Língua Portuguesa); o VOP (Vocabulário Ortográfico do Português – Portugal); e o VO-TL (Vocabulário Ortográfico de Timor-Leste). Disponibiliza, de forma organizada e homogénea, os dados de cada vocabulário nacional, bem como as particularidades de cada variedade, diretamente decorrentes, ou não, das bases do AO90. Presentemente, o VOC não alberga os vocabulários nacionais de Angola e Guiné-Bissau, por

---

27 Os dados apresentados neste subponto estão disponíveis em <http://www.portaldalinguaportuguesa.org/?action=acordo-historia> (página consultada em maio de 2016), com organização cronológica mais detalhada.

28 O Segundo Protocolo Modificativo ao Acordo Ortográfico da Língua Portuguesa, assinado pelos então oito membros da CPLP, além de autorizar a adesão de Timor-Leste à comunidade, definiu a entrada em vigor do AO90 com o depósito dos instrumentos de ratificação por três países signatários. À data da redação deste trabalho, o AO90 encontra-se ratificado por Brasil, Cabo Verde, Guiné-Bissau, Portugal, São Tomé e Príncipe e Timor-Leste.



constrangimentos que condicionaram um consenso relativamente a uma norma ortográfica comum a vários estados soberanos, entendida como própria por cada uma das nações signatárias de um acordo internacional. Questões identitárias podem ser apontadas como motivo para a falta de acordo, bem como questões técnicas mal esclarecidas ou interpretadas de forma não consensual. Mais complicadas do que estas, em casos particulares existem questões de âmbito político e militar que não contribuem para uma discussão produtiva com vista à elaboração de recursos deste tipo. A instabilidade de um país, seja de que tipo for, serve de bloqueio à implementação de diretrizes relacionadas com política linguística, colocando-as, justificadamente, em segundo plano.

O VOC é a soma das partes que o constituem mas, mais do que isso, é a representação de uma “língua portuguesa pluricêntrica”, uma plataforma comum que se propõe a ser partilhada por todos os falantes de língua portuguesa, construída tendo presentes as diferenças entre eles.

A base de dados lexicais assenta no modelo OSLIN, tecnologia desenvolvida pelo ex-ILTEC (Instituto de Linguística Teórica e Computacional) que segue a interface do Portal da Língua Portuguesa, incorporando em mais de 98% código-fonte desenvolvido no âmbito desta interface. O VOC é desenvolvido pelo IILP, em parceria com o agora CELGA-ILTEC (Centro de Estudos de Linguística Geral e Aplicada - Instituto de Linguística Teórica e Computacional) e com o NILC (Núcleo Interinstitucional de Linguística Computacional). A plataforma segue princípios metodológicos coerentes e os seus dados resultam de um entendimento comum para a aplicação plena do AO90 por parte de especialistas de cada estado-membro indicados para o efeito. É realizada uma homogeneização e integração progressiva dos vocabulários ortográficos nacionais elaborados por parceiros nacionais indicados institucionalmente por cada um dos governos, sendo posteriormente acompanhada a evolução da base de dados.<sup>29</sup>

---

29 Veja-se, a este respeito, Almeida et al. (2013) e Ferreira e Correia (2014).



## 4. O VOCABULÁRIO TOPONÍMICO

Incorporados no VOC, têm vindo a ser desenvolvidos dois recursos que almejam alcançar e tratar áreas críticas da ortografia bem identificadas: o Dicionário de Formas Não Adaptadas (DFNA) e o Vocabulário Toponímico (VT).

O primeiro atesta palavras que são tradicionalmente designadas por «estrangeirismos» ou «empréstimos» (estas designações remetem para algo «externo à língua portuguesa», o que, em rigor, não pode ser assumido como válido). Na realidade, estas palavras correspondem a vocábulos integrados na língua mas que simplesmente não possuem forma ortográfica adaptada. Pela interpretação desta realidade, o DFNA atesta «formas não adaptadas», e não «estrangeirismos» ou «empréstimos». As formas não adaptadas podem ter um sinónimo perfeitamente em conformidade com as regras de escrita da língua portuguesa, como acontece com «casting», que tem como sinónimo «audição», ou serem formas exclusivas, como ocorre com «cheddar». Podem igualmente ser formas não preferenciais relativamente a uma forma adaptada, mas ainda em uso, como acontece com «pivot» e «pivô», respetivamente.

O VT, a par do DFNA, procura colmatar lacunas relativas à adequada atestação de formas ortográficas pertencentes a um subconjunto lexical específico bem delimitado. Nele estão atestados os nomes de todos os países do mundo reconhecidos pela Organização das Nações Unidas (ONU), suas capitais, a toponímia maior de Brasil, Cabo Verde, Moçambique, Portugal, São Tomé e Príncipe e Timor-Leste. Além destes, alguma toponímia menor e topónimos com relevância atribuída pelo uso ou por questões de proximidade histórico-cultural relativamente aos países da CPLP incorporarão a primeira versão do recurso. Todos os topónimos atestados têm indicação das propriedades formais, nomeadamente a divisão silábica e a sílaba tónica. Nos casos em que for possível aferi-la, é incluída informação relativamente à obrigatoriedade, facultatividade ou impossibilidade de emprego de artigo aquando da nomeação do topónimo em contexto. É também dada informação explícita relativa à variação, nos casos em que ela exista. Sempre que possível, está estabelecida uma relação entre o topónimo e o(s) gentílico(s) correspondente(s), derivacionais ou não. A norma ortográfica seguida é a vigente para a língua portuguesa (AO90).

### 4.1. Nomeação do recurso

O Vocabulário Toponímico, ou VT, é assim designado por se tratar de uma sistema lexical de especialidade que inclui dados gerais para a identificação inequívoca de topónimos, entre eles informação semântica explícita de tipo relacional (relação hierárquica entre topónimos e seu correspondente enquadramento numa determinada área geográfica), bem como a classe de palavra implícita dos topónimos (sempre nomes próprios). O VT é um recurso digital, portanto, logicamente, não está organizado da mesma forma que um recurso semelhante em papel. Não obstante, é possível aceder a uma lista com todos os topónimos por ordem alfabética. Mais importante, a rede de topónimos disponibilizada constitui real inovação para uma língua pluricêntrica sob a mesma norma ortográfica, na medida em que se apresenta

como recurso de tipo vocabularístico único para a toponímia em língua portuguesa, atualizável e passível de ser aprimorado à medida que a realidade se vai alterando.

Em suma, os dados compilados no VT, sendo efetivamente de tipo vocabularístico, possuem, adicionalmente, informação semântica relacional entre topónimos, assim como entre gentílicos (não atestados diretamente no VT) e os referentes geográficos que lhes dão origem. Também por não possuir informação semântica tradicionalmente incluída em dicionários, nomeadamente a de natureza etimológica, adotou-se a designação «Vocabulário Toponímico».

#### 4.1.1. Livre acesso a dados toponímicos

O VT possibilita aos seus consulentes esclarecer dúvidas de diferentes tipos. Tal como ocorre na consulta a um dicionário geral da língua, a pesquisa pode estar relacionada com a verificação da existência de uma palavra (neste caso, de um topónimo); com a necessidade de esclarecer de que forma deve ser grafada uma palavra ou de aferir quais as suas propriedades formais; com a procura por informação semântica (no caso do VT, do posicionamento hierárquico do topónimo relativamente a outros, ou da relação entre topónimos e gentílicos do léxico geral). Independentemente da motivação que esteja na origem da consulta, toda a informação disponibilizada no VT é de acesso gratuito. Pelo enquadramento institucional e político do próprio VOC, o posicionamento do VT não poderia ser outro, sendo imperativo o seu explícito distanciamento relativamente a interesses económicos ou de âmbito privado. Desta forma, vê-se reforçado o interesse comum dos falantes de língua portuguesa (e não só) e o direito à livre informação relativamente a esta área do conhecimento.

## 4.2. Padronização: mais-valias

A adequada padronização da toponímia é um objetivo almejado por qualquer sistema linguístico do mundo. Serve, enquanto recurso de e para as populações, como elemento reforçativo da assunção de cidadania e, de um ponto de vista de aplicação efetiva em sistemas de ordenação geográfica, como recurso cujas linhas gerais se tornam indispensáveis para uma correta referenciação do território. Se a sua implementação não ocorre em algumas línguas, tal deve-se à falta de coordenação entre as entidades competentes ou a uma errada secundarização da temática. Ao nível da Organização das Nações Unidas, tem vindo a ser desenvolvido trabalho de há umas décadas a esta parte no sentido da padronização da toponímia de vários países, trabalho que deve ser seguido com atenção por parte dos países da CPLP. Esta análise é desenvolvida no ponto 4.7 desta dissertação.

Com o intuito de padronizar a toponímia em língua portuguesa, o recurso a bases de dados toponímicas é fundamental, permitindo a consolidação e divulgação de topónimos únicos no espaço da CPLP de forma homogénea. A mesma premissa se aplica às diferentes línguas do mundo, uma grande parte das quais ainda sem adequados recursos humanos e sem aposta tecnológica dedicada ao desenvolvimento deste trabalho.

### 4.3. Critérios lexicais: seleção de topónimos

Os topónimos atestados no VT foram selecionados tendo em conta os dados toponímicos dos institutos nacionais de estatística, bem como dos CTT, no caso de Portugal. Todos os nomes de lugares com relevância administrativa grafados em língua portuguesa e pertencentes à toponímia maior do espaço da CPLP foram *a priori* considerados topónimos válidos e obrigatórios para integração no VT. Adicionalmente, os topónimos referentes a países e suas capitais foram integrados no recurso, tendo presentes os dados disponibilizados pelo IBGE (Brasil) e pela União Europeia (Código de Redação Interinstitucional). Gentílicos derivados destes topónimos, quando atestados em fontes de referência e integrados no léxico comum, têm marcação de relação derivacional com os topónimos que formam a nomenclatura do VT. Outros nomes de lugares localizados fora do espaço da CPLP que não países ou capitais de país foram – pontualmente, pela pertinência histórica, proximidade cultural ou relevância decorrente do uso comum e generalizado – integrados no VT, igualmente com ligação aos gentílicos deles derivados já atestados no léxico geral (e.g. Macau e macaense). Adicionalmente, quando disponíveis, nomes de lugares sem relevância administrativa no espaço da CPLP foram incluídos (aqui se incluem os distritos do Brasil, por exemplo). O processo de verificação e atestação destes topónimos está ainda numa primeira fase, não estando, à data, tão sistematicamente desenvolvido.

#### 4.3.1. Toponímia maior e toponímia menor

Ainda que seja correto assumir que a toponímia estuda os «topónimos», ou os «nomes de lugares», importa dar conta da possível e pertinente distinção entre «toponímia maior» e «toponímia menor», bem como da útil associação da designação «nome geográfico» a um subconjunto da toponímia menor.

Como se infere da análise às referências dicionarísticas citadas no ponto 2., os nomes de lugares sem relevância administrativa (toponímia menor) são comumente designados apenas por «topónimos». No entanto, importa ter presente a fronteira entre topónimos «maiores» e topónimos «menores», que possibilita, por exemplo, a identificação e o tratamento sistemáticos de um conjunto de referentes mais estáveis – a toponímia maior. O VT, enquanto primeira versão de um recurso de política linguística transnacional, tem esta fronteira terminológica presente.

Garnelo (2001:71) refere a divisão entre toponímia maior e toponímia menor, embora não a subscreva, por considerar pouco claros os motivos que estão na sua base. No entanto, o que é dito pelo autor acaba por reforçar a potencial utilidade do reconhecimento das vantagens decorrentes desta partição na feitura de um recurso como o VT, que numa primeira fase pretende ser o mais abrangente possível sem descuidar a solidez dos dados disponibilizados, não esquecendo o seu carácter normalizador e a sua motivação também política. Com esse intuito, o VT tem por obrigatória a atestação das formas mais estáveis, geralmente mais bem documentadas e com menor margem para erros provenientes das fontes escrutinadas (a toponímia maior), construindo de raiz um recurso suficientemente sólido, ao ponto de, no futuro, ser possível dissipar a

fronteira, agora útil, entre toponímia maior e toponímia menor, pela inclusão sistemática da toponímia menor em língua portuguesa no espaço da CPLP.

Segundo uma distinção universalmente aceite, a toponímia divide-se em toponímia maior e menor; porém, é difícil precisar quais são os critérios distintivos entre uma e outra. A toponímia maior ocupa-se, para uns, dos nomes dos lugares de especial importância, como os nomes das povoações [...]. A toponímia menor ocupar-se-ia dos nomes secundários, ou dos nomes de espaços mais reduzidos. Em minha opinião, é uma distinção absolutamente artificial e desnecessária. É verdade que se podem estabelecer algumas diferenças entre os dois grupos. A toponímia maior, ao designar lugares maiores e mais importantes, é mais estável. Além disso, está geralmente mais bem documentada e é possível rastrear mais facilmente as suas origens. Por outro lado, é mais arbitrária; o nome pode responder a motivações muito diversificadas, algumas vezes totalmente alheias ao lugar que designa. A toponímia menor é mais precisa, nasce de uma motivação direta entre o objeto e o nome, está menos documentada e é mais variável, porque os habitantes mudam o nome por causa de múltiplas causas; também é mais instável, e os nomes perdem-se, substituídos por outros novos, com relativa frequência. (Garnelo 2001:71) [tradução própria]

Pelo dito anteriormente e pela citação imediatamente acima, os nomes de lugares sem relevância administrativa,<sup>30</sup> ou, se preferirmos, hierarquicamente abaixo da unidade com relevância administrativa mínima, são então assumidos no VT como parte da toponímia menor, sem atestação sistemática no recurso, sendo ainda residuais. Veredas, ruas, largos, praças, avenidas ou logradouros, enquanto «nomes geográficos», pertencem igualmente ao ramo da toponímia menor. A designação «menor» procede apenas da dimensão geográfica dos referentes, sempre parte constituinte de unidades maiores, não existindo qualquer intenção de os minorizar ou aviltar, ou a quem por eles nutra especial afeto. Por não ser a toponímia menor, que inclui os nomes geográficos, trabalhada transversalmente na primeira versão do VT, não se desenvolve, então, a sua aturada análise nesta dissertação.

Os topónimos e os nomes geográficos têm, ainda assim, características partilhadas, entre as quais convém destacar a de mormente provirem ou de um antropónimo ou de um nome comum, isto é, de possuírem um étimo que sincronicamente poderá ser transparente, seja por se referir a acidentes geográficos (e.g. Rio Tinto), a hierónimos (e.g. São Paulo) ou à predominância local de uma espécie animal ou vegetal (e.g. Açores), entre outros. Devem também ser distinguidos dois significados para «lugar»: um, mais abrangente, que corresponde a qualquer unidade geográfica, independentemente da sua dimensão e representatividade político-administrativa; outro, mais específico, que corresponde a unidades geográficas menores, em Portugal abaixo das freguesias mas não necessariamente circunscritas a uma só freguesia, conforme é descrito pelo INE na sua página oficial: “o lugar consiste numa delimitação territorial, definida no âmbito das operações censitárias, que corresponde a um aglomerado populacional com dez ou mais alojamentos destinados à habitação de pessoas e com uma designação própria, independentemente de pertencer a uma ou

---

30 Ver ponto 4.3.2. desta dissertação.

mais freguesias.” A definição do INE corresponde então a nomes geográficos que não estão por agora, na sua esmagadora maioria, atestados no VT. Nesta dissertação, salvo referência explícita, «lugar» tem significado mais abrangente.

#### 4.3.2. Relevância administrativa

A relevância administrativa que um lugar pode ou não ter é contida pelas fronteiras artificiais que o definem, delimitada por uma área geográfica identificável escorreitamente pelas populações. Trata-se de uma dimensão política considerada relevante na exata medida em que autoriza o lugar enquanto unidade de referência junto de uma população. A relevância administrativa está então associada a uma unidade geográfica politicamente autónoma ou que possui determinadas atribuições e competências de alguma forma independentes do poder central, ainda que com posição hierárquica menor relativamente a este. Em Portugal, por exemplo, a unidade mínima com relevância administrativa é a freguesia, designação atual para as antigas paróquias civis. Mais do que dar conta da alteração ao nome do conjunto destas unidades, importa referir que as «unidades mínimas com relevância administrativa» não têm caráter permanente, ou seja, pode dar-se o caso de uma unidade com relevância administrativa hoje não a possuir amanhã. Veja-se a Lei de 20 de outubro de 1840 e o Código Administrativo de 1842: retiraram, ainda que temporariamente, a relevância administrativa às freguesias, atribuindo-a somente aos distritos e municípios. Até à aplicação do Código Administrativo de 1878, que restituiu algumas diretrizes do Código Administrativo de 1836, nomeadamente a reintegração das freguesias (então paróquias) como unidades com relevância administrativa (Art. 1.º), foram os municípios as unidades menores consideradas. Mesmo nos períodos em que as freguesias possuam relevância administrativa, a existência destas unidades não está assegurada: entre processos de atribuição de competências ao nível da freguesia, muitas desaparecem, outras são criadas. A título de exemplo, a mais recente reorganização administrativa em Portugal reduziu o número de freguesias em Lisboa para cerca de metade, de 53 para 24.<sup>31</sup>

As competências de cada freguesia podem variar mas estão normalmente bem balizadas e pressupõem uma política de proximidade e contacto direto com as populações locais. Continuando com o exemplo de Lisboa, embora os espaços, equipamentos e vias estejam sob a responsabilidade da câmara municipal, as juntas de freguesia tratam de questões mais circunscritas, como sejam a habitação ou a intervenção comunitária, os licenciamentos de proximidade (e.g. “licenciamento para atividades ruidosas”) ou a conservação e reparação de chafarizes.<sup>32</sup>

Em Portugal, além das freguesias, existem enquanto unidades com relevância administrativa os municípios e os distritos, bem como os níveis NUTS I, II e III (ver ponto 5.2.4.). No VT, como referido no ponto 2. desta dissertação, foi usada a noção de «relevância administrativa» associada a topónimos enquanto principal propriedade para o estabelecimento da nomenclatura do VT.

---

31 Ver Deliberação n.º 6/AML/2014, de 22 de janeiro de 2014 (página consultada em fevereiro de 2016).

32 Veja-se <http://www.cm-lisboa.pt/servicos/por-temas/reforma-administrativa-de-lisboa/competencias-das-juntas-de-freguesia> (página consultada em fevereiro de 2016).

## 4.4. Filiação histórica na área da filologia

O estudo da toponímia tem, tradicionalmente, particular foco na etimologia. Um dos autores mais autorizados neste ramo em Portugal e, de uma forma geral, no que à onomástica diz respeito, é Rebelo Gonçalves. Muitos outros se lhe seguiram, entre os quais José Pedro Machado. Aturado trabalho foi realizado nos últimos séculos com o intuito de aprimorar o tratamento e a sistematização de informação, por vezes difusa, sobre a origem e o percurso histórico da toponímia maior, mas os trabalhos desenvolvidos até hoje esbarram na dificuldade em atestar de forma coesa, com fontes minimamente fidedignas, toda a toponímia em língua portuguesa.

A abordagem do VT afasta-se da filiação histórica da toponímia na área da filologia, focando o estudo sincrónico dos topónimos e os conceitos de norma ortográfica, uso e padronização toponímica.

## 4.5. Tradição normativa para a toponímia

Os topónimos, como as restantes palavras que constituem o léxico de uma língua, respeitam as normas ortográficas em vigor para o idioma em que são grafados. Enquanto ramo da onomástica, no entanto, a toponímia possui algumas especificidades, sendo abrangida por um subconjunto de regras ortográficas. Desde 1911, os textos legais para a língua portuguesa dão conta disso mesmo, enquadrando os topónimos numa classe de palavras sob condição particular.

O primeiro documento normativo com força de lei para a língua portuguesa data de 1911, ano em que o Formulário Ortográfico (FO11) entrou legalmente em vigor. O FO11, baseado no Vocabulário Ortográfico e Ortoépico da Língua Portuguesa (1909), da autoria de Gonçalves Viana, teve como principal intuito minimizar uma deriva ortográfica que se vivia até então. No que diz respeito à toponímia, no FO11 apenas é referido o caso da adaptação de «Antwerp» para «Antuérpia» e a possibilidade de utilização das letras <k>, <w> e <y> em “vocábulos derivados de nomes estrangeiros”, autorizando uma grafia peculiar em determinados gentílicos.<sup>33</sup>

O Acordo de 1931, por divergências entre o Brasil e Portugal, acabaria por não ser posto em prática. No entanto, nas Conclusões Complementares do Acordo de 1931 (anexo do AO45) acrescenta-se referência aos topónimos terminados em consoantes que em português estão normalmente associadas à posição de ataque de sílaba<sup>34</sup> (i.e. <b>, <c>, <d>, <g> e <t>).

É sublinhada a desejabilidade do emprego de formas vernáculas na grafia de topónimos estrangeiros, relacionando esta prática não apenas com as formas já consolidadas na língua mas também com o “uso comum”: “[deve proceder-se à] substituição de topónimos de línguas estrangeiras por formas vernáculas equivalentes, sempre que estas sejam antigas na Língua ou entrem no uso corrente.”<sup>35</sup>

33 Formulário Ortográfico de 1911, Base I - Letras k, w, y - ponto I.

34 Primeira consoante ou duas primeiras consoantes contíguas de uma sílaba.

35 Conclusões Complementares do Acordo de 1931, Segunda Parte - Ponto 51.



No Formulário Ortográfico de 1943 (FO43), oficializado e aplicado somente no Brasil, não é de somenos a prescrição de uma forma preferencial para grafar os topónimos de origem estrangeira, sendo instituída a transcrição segundo as normas constantes das Resoluções da Conferência de Geografia de 1926,<sup>36</sup> desde que em harmonia com o próprio FO43, no caso de se verificar a inexistência de formas vernáculas.<sup>37</sup> Neste instrumento, é pela primeira vez referido o vocabulário comum enquanto ferramenta para validação de compostos e derivados de topónimos.

O Acordo Ortográfico de 1945 (AO45), nas suas Bases Analíticas, desenvolve o que é prescrito na Segunda Parte - Ponto 8 das Conclusões Complementares do Acordo de 1931, sendo feita referência à oralidade e à “tradição bíblica”, com exemplos.<sup>38</sup>

#### 4.5.1. Acordo Ortográfico de 1990

O Acordo Ortográfico de 1990 (AO90), redigido tendo por base o texto do AO45, logo, utilizando várias passagens deste praticamente inalteradas, na sua “Base I: do alfabeto e dos nomes próprios estrangeiros e seus derivados”, segue a tradição de especificar o uso das letras <k>, <w> e <y> na onomástica: “As letras *k*, *w* e *y* usam-se [...] em topónimos/topônimos originários de outras línguas e seus derivados: *Kwanza*, *Kuwait*, *kuwaitiano*; *Malawi*, *malawiano*.”<sup>39</sup>

É igualmente replicado o Ponto 8 do AO45 (com pequenas alterações ao nível de construção frásica sem necessidade de referência), que remete para a ocorrência das consoantes finais <b>, <c>, <d>, <g> ou <h> e o uso a dar-lhes em língua portuguesa.<sup>40</sup>

#### 4.5.2. Atualidade e aplicação de regras específicas aos topónimos

Os topónimos, enquanto nomes próprios, obedecem às regras gerais de ortografia, com propriedades particulares amplamente descritas na literatura. Ainda assim, mais do que a mera obrigatoriedade de emprego de letra maiúscula inicial em quaisquer palavras plenas que o constituam – aqui se incluem os topónimos compostos –, existem regras de hifenização, acentuação e respeitantes a sequências consonânticas próprias dos topónimos.

Também os apóstrofes são diacríticos possíveis na grafia de topónimos (e.g. Paço d'Arcos) mas, em conformidade com os critérios adotados para as restantes formas lexicais do VOC, foram preteridos, assumindo-se sempre graficamente no VT a vogal que pretendem elidir (e.g. Paço de Arcos). Assim, o

---

36 Na Conferência de Geografia de 1926 foi iniciado o trabalho de padronização dos topónimos brasileiros. Diziam as resoluções: “Que se grafe com ”j”, e não com “g”, o fonema “jê” dos nomes de origem indígena ou africana; Que se grafem com “que” os finais dos nomes de origem indígenas ou africana em que atualmente se emprega ora com “c”, ora com “k”, ora com “ck”; Que se grafe com “x” o som chiante dos nomes de origem indígena ou africana; Que se substitua nos nomes de origem indígena ou africana o “y” pelo “i”; Que se escreva “z”, e não “s” entre vogais, nos sobrenomes de origem indígena ou africana.”

37 Formulário Ortográfico de 1943, Base XI - Nomes Próprios - Ponto 41.

38 Bases Analíticas do Acordo Ortográfico de 1945, Documento n.º 2, Ponto 8.

39 Acordo Ortográfico de 1990, Base I, ponto 2 b).

40 Acordo Ortográfico de 1990, Base I, ponto 5.

emprego de apóstrofo como forma de “cindir graficamente uma contração ou aglutinação vocabular”, conforme descrito na Base XVIII do AO90, não deixa de ser válido nos casos em que a tradição o consagre, ainda que sem correspondência atestada no VT. A possibilidade de uso de apóstrofo por forma a marcar uma elisão de vogal «fraca» é conferida por critérios normalmente relacionados com questões de estilo, não raras vezes refletidos na tradição mas que, necessariamente, tornam difícil aferir grafias simples e homogêneas de topónimos, causando variação indesejável, motivo pelo qual não são incluídos apóstrofos nesta obra normalizadora. Tendo o texto do AO90 como instrumento para a aplicação de critérios ortográficos aos topónimos, é útil analisá-lo quanto a cada uma das especificidades trabalhadas.

#### 4.5.2.1. Hifenização

O AO90 faz referência aos contextos em que deve grafar-se um topónimo com recurso ao hífen, dando exemplos e clarificando (também com exemplos) os casos em que o seu uso é rejeitado:

“Emprega-se o hífen nos topónimos/topônimos compostos iniciados pelos adjetivos *grã*, *grão* ou por forma verbal ou cujos elementos estejam ligados por artigo: *Grã-Bretanha*, *Grão-Pará*; *Abre-Campo*; *Passa-Quatro*, *Quebra-Costas*, *Quebra-Dentes*, *Traga-Mouros*, *Trinca-Fortes*; *Albergaria-a-Velha*, *Baía de Todos-os-Santos*, *Entre-os-Rios*, *Montemor-o-Novo*, *Trás-os-Montes*. Obs.: Os outros topónimos/topônimos compostos escrevem-se com os elementos separados, sem hífen: *América do Sul*, *Belo Horizonte*, *Cabo Verde*, *Castelo Branco*, *Freixo de Espada à Cinta*, etc. O topónimo/topônimo *Guiné-Bissau* é, contudo, uma exceção consagrada pelo uso.”<sup>41</sup>

Assim, para a língua portuguesa esta base introduz três (de quatro) contextos em que deve empregar-se hífen em topónimos:

- a) topónimos iniciados por *Grã-* ou *Grão-*;
- b) topónimos iniciados por forma verbal;
- c) topónimos cujos elementos estejam ligados por artigo.

Assumido como exceção a estas quatro regras, o topónimo «Guiné-Bissau» deve ser igualmente grafado com hífen, com o pressuposto de tal grafia estar consagrada pelo uso. À semelhança do carácter excecional atribuído ao topónimo «Guiné-Bissau», também o topónimo «Timor-Leste» é grafado com hífen, segundo decisão tomada pelo governo leste-timorense. A exceção para «Timor-Leste» não está integrada no texto legal, uma vez que a plena independência do país ocorreu no dia 20 de maio de 2002, doze anos após a assinatura do acordo pelos países então envolvidos, tendo Timor-Leste aderido à CPLP nesse mesmo ano da década de 2000, assinando posteriormente o Segundo Protocolo Modificativo ao Acordo Ortográfico da Língua Portuguesa, em 2004.

---

41 Acordo Ortográfico de 1990, Base XV, ponto 2.

#### 4.5.2.1.1. Tupi-guarani

O tupi-guarani é uma família linguística de grande relevância – especificamente no que respeita à riqueza e património linguísticos do Brasil e, de uma forma mais abrangente, na relação histórica com a língua portuguesa, nesta tendo deixado marcas indeléveis. As línguas da família tupi-guarani assumiram, portanto, desde o período colonial português, papel de línguas de contacto na relação com a língua portuguesa. Mello (2000), com o auxílio de uma referência geográfica concisa, diz-nos isso mesmo:

[Da família tupi-guarani fazem parte] aproximadamente quarenta línguas fortemente relacionadas, com uma distribuição geográfica bastante ampla na América do Sul. Seus extremos atuais são o litoral do Brasil (leste), as margens do Amazonas na fronteira Brasil-Peru (oeste), sul da Guiana Francesa (norte) e sul do Brasil, Paraguai e norte da Argentina (sul). [...] Foi amplamente documentada no período colonial da história do Brasil, com finalidades mercantis e religiosas, e nesta época era denominada Língua Brasília, ou “a língua mais falada da costa do Brasil”. (Mello 2000:11)

De forma previsível, o tupi-guarani contribuiu em grande medida para a constituição de muitos novos vocábulos da língua portuguesa, entre eles, como não podia deixar de ser, topónimos brasileiros que não possuíam atestação em línguas que não as locais. Da tradição de topónimos grafados em língua portuguesa provenientes do tupi-guarani ou que dele incorporaram um constituinte morfológico, chegamos à formulação, no AO90, de uma regra específica que obriga à adição de um quarto – e último – caso em que deve grafar-se um topónimo com recurso ao hífen:

“Nas formações por sufixação apenas se emprega o hífen nos vocábulos terminados por sufixos de origem tupi-guarani que representam formas adjetivas, como *açu*, *guaçu* e *mirim*, quando o primeiro elemento acaba em vogal acentuada graficamente ou quando a pronúncia exige a distinção gráfica dos dois elementos: [...] *Ceará-Mirim*.”<sup>42</sup>

d) topónimos terminados em sufixos de origem tupi-guarani, quando o primeiro elemento termina em vogal acentuada graficamente ou quando a pronúncia a tal obriga.

Quer isto dizer que todos os topónimos terminados em sufixo de origem tupi-guarani com função adjetival original devem ser hifenizados apenas quando o primeiro elemento termina em vogal acentuada graficamente, ou tendo por base a pronúncia não desviante do topónimo (e.g. Sapucaú-Mirim). Esta regra incluirá primeiros elementos com a antepenúltima vogal acentuada graficamente (proparoxítonos), que, pelo sinal gráfico obrigatório, teriam diferente pronúncia caso esse acento agudo fosse omitido por um processo de aglutinação. A distinção gráfica poderia ser feita pela simples separação dos dois elementos mas, seguindo a letra do AO90, “emprega[-se] o hífen nos vocábulos terminados por sufixos de origem tupi-guarani que

---

42 Acordo Ortográfico de 1990, Base XVI, ponto 3.

representam formas adjetivas, como *açu*, *guaçu* e *mirim*, (...) quando a pronúncia exige a distinção gráfica dos dois elementos”. A pronúncia de uma forma proparoxítona exige, pois, a distinção gráfica dos dois elementos, logo, deve proceder-se à hifenização. Nos restantes casos teremos ora formas aglutinadas (e.g. Itapemirim), ora formas independentes (e.g. Itapecuru Mirim).

Relativamente aos gentílicos provenientes de topónimos com constituintes de origem tupi-guarani, aplica-se a regra geral: são hifenizados todos os gentílicos compostos, independentemente do topónimo que lhes dá origem (e.g. Tomé-Açu → tomé-açuense; Conceição do Lago Açu → lago-açuense).

#### 4.5.2.2. Acentuação

A acentuação gráfica de topónimos em língua portuguesa não é, por princípio, problemática. Existe variação entre o emprego de acento circunflexo (no Brasil) ou agudo (nos restantes países da CPLP) em topónimos proparoxítonos cuja vogal tónica <e> ou <o> é imediatamente seguida de consoante nasal <m> ou <n> (e.g. Polônia/Polónia), mas a regra de acentuação é agora única, sendo apenas distinta a sua aplicação consoante a variedade em uso.<sup>43</sup>

São também pontuais os casos em que um topónimo possui variação entre o emprego e a ausência de diacrítico para marcação de tonicidade. Após a entrada em vigor do AO90, este esporádico tipo de variação – verificável, por norma, dentro de uma mesma variedade, e atestado em diferentes placas toponímicas para um mesmo lugar – será progressivamente suprimido, tendendo a desaparecer. Na freguesia de Armação de Pera (município de Silves), por exemplo, a variação existia em 2009 e está ainda por resolver;<sup>44</sup> o mesmo ocorre na antiga freguesia de Pero Pinheiro (hoje freguesia de Almargem do Bispo, Pero Pinheiro e Montelavar, do município de Sintra).<sup>45</sup> Além destes, existem casos gritantes em que dificilmente se pode compreender a efetiva colocação de placas toponímicas com ortografia desviante, nomeadamente no que diz respeito à acentuação gráfica. Crastos, lugar da freguesia de Porches, município de Lagoa, é disto exemplo, possuindo no mínimo uma placa toponímica em que, desde pelo menos 2009, subsiste um erro ortográfico grosseiro, estando o nome do lugar atestado como «Crástos».<sup>46</sup>

De sublinhar os topónimos com ditongo decrescente de núcleo tónico, que deixam, todos, de ter diacrítico.<sup>47</sup> Um dos exemplos com mais visibilidade em Portugal, ainda que o referente não possua relevância administrativa, é o do nome «Troia», que identifica um lugar da freguesia do Carvalhal, no município de Grândola. O nome «Tróia», forma pré-AO90, passa a grafar-se «Troia». Mesmo que a alteração tenha já sido amplamente divulgada e discutida, e que haja diversas unidades hoteleiras locais cujos nomes nunca incluíram acento gráfico,<sup>48</sup> as placas toponímicas de Troia, em janeiro de 2015, não haviam ainda sido atualizadas (Figura F, em Anexos). Em 2009 existia variação nas placas toponímicas de Troia: numa placa

43 Acordo Ortográfico de 1990, Base XI, ponto 3.

44 Ver Figura A e Figura B, em Anexos.

45 Ver Figura G e Figura H, em Anexos.

46 Ver Figura I e Figura J, em Anexos.

47 Acordo Ortográfico de 1990, Base IX, ponto 3.

48 Este facto não é determinante para a atualização das placas toponímicas em Troia ou em qualquer outro lugar mas, sendo parte integrante da paisagem linguística local, pode reforçar a perceção da necessidade de atualização.

que em agosto desse ano assinalava a fronteira do lugar, o nome era grafado sem acento (Figura C, em Anexos) e noutra, também em agosto de 2009, o nome surgia grafado com acento (Figura E, em Anexos). Aparentemente, a variação, pelo menos no que respeita às duas placas aqui observadas, foi minorizada, pela substituição<sup>49</sup> da placa aqui identificada na Figura C. Mas ocorre que, na realidade, o nome aí atestado passou a estar grafado com acento (Figura D, em Anexos), não tendo a atualização sido feita de forma adequada, depreende-se que por lapso. A simples consulta dos critérios do VT, mesmo não estando ainda no recurso atestado este nome de lugar sem relevância administrativa, resolveria quaisquer dúvidas que tenham surgido na altura da atualização das placas do lugar, com mais-valias decorrentes de uma atualizada paisagem linguística local, evitando custos económicos futuros decorrentes de nova substituição das placas na península.

Existem ainda casos em que é necessário, em termos de acentuação gráfica, adaptar topónimos cujo referente geográfico esteja localizado no espaço da CPLP mas que são tradicionalmente grafados em línguas que não a portuguesa, com regras ortográficas necessariamente distintas das do português. Um exemplo desta realidade será a adaptação para o português de alguns topónimos de Moçambique com acento grave. Nas línguas nacionais moçambicanas que não a portuguesa, assume-se nesta dissertação que, de uma forma geral, o acento grave nas palavras terminadas em *-uè* serve para marcar a altura da vogal (átona) nos casos em que esta terminação é antecedida de uma vogal com acento agudo, e não para marcar a sua tonicidade (a palavra será proparoxítone ou paroxítone com ditongo decrescente). O acento grave na última vogal dará igualmente indicação de que a vogal imediatamente anterior é pronunciada (exemplo *i*), abaixo) e faz parte de uma outra sílaba (exemplo *ii*), abaixo) – existirá um hiato –, sílaba essa que poderá incluir ou não um ditongo. Quando a terminação não é antecedida de vogal com acento agudo, pode colocar-se a hipótese de o acento grave servir para marcar a tonicidade da palavra (palavra oxítone) e/ou para dar conta da prolação da vogal imediatamente anterior <u>. Assim:

i) *Súbwè* → *Súbue* [‘subuE]

ii) *Muxunguè* → *Muxungué* [muSũgu‘E]

#### 4.5.2.3. Sequências consonânticas

As sequências consonânticas *-cc-*, *-cç-*, *-ct-*, *-pc-*, *-pç-* e *-pt-*, que tanta discussão suscitaram aquando da aprovação do AO90 – ainda que de forma mais ruidosa apenas no período imediatamente anterior à sua entrada em vigor – não são problemáticas quando analisamos a sua ocorrência em topónimos. Existem, no entanto, casos pontuais em que estas sequências se mantêm, podendo ou não corresponder a duas consoantes oclusivas na oralidade. Tal apenas se aplica, ainda assim, no caso de os topónimos referidos provirem de um antropónimo, mantendo a forma gráfica deste mas sendo reinterpretados enquanto topónimos. A título

---

49 Não foi possível aferir a data exata de substituição da placa (as imagens disponíveis são de dezembro de 2014 e de janeiro de 2015), que terá ocorrido entre 2009 e 2014.

exemplificativo, podemos referir o município brasileiro Januário Cicco, no estado do Rio Grande do Norte; ou a freguesia cabo-verdiana de São João Baptista, no concelho de Porto Novo.

Já o caso de «Egito» é aquele em que a supressão de uma consoante mais celeuma suscitou. Assumiu-se, aquando da redação do AO90, que a prolação-padrão do topónimo não incluiria a consoante [p], logo, que não deve ser grafada a letra <p> no topónimo. Acontece que a exclusividade desta prolação não é consensual, existindo argumentos a favor e contra tal restrição ortográfica. O VT, quanto a este caso específico, limita-se a atestar a forma indicada no texto legal. A restrição para o topónimo, útil do ponto de vista funcional pela ausência, neste caso, de variação intravarietal, não invalida, claro está, a grafia discrepante do gentílico associado a este topónimo, mas que na realidade é reflexo da prolação única, transversal a todas as variedades da língua portuguesa, do gentílico com consoante [p], logo, sempre grafado «egípcio». Note-se que o par topónimo-gentílico «Egito»/«egípcio» não é caso único de discrepância entre o radical de um e outro elementos. Embora, pela transparência desejável, estes pares devam possuir o maior grau possível de similaridade, tal nem sempre é efetivado, mormente por questões históricas. Veja-se, apenas para os nomes de países, os pares «Chipre»/«cipriota», do grego *kýpros* e do latim *Cyprus* ou *Cypros*; «Equador»/«equatoriano», em que o gentílico mantém o <t> do latim *aequator*; «Madagáscar, Madagascar»/«malgaxe», em que a forma gentílica parte ainda do antigo topónimo «República Malgaxe», criado em 1960, no seguimento da independência do país; «Mónaco»/«monegasco», cujo gentílico é em Machado (2003:1013) descrito como sendo “do fr. *Monégasque*, certamente formado a partir do lat. *Monoecus* (ver *Moneco*)<sup>50</sup> por via culta”; ou «Seri Lanca»/«cingalês», cujo gentílico deriva da antiga designação para o país em língua portuguesa, «Ceilão», pela etnia cingalesa.<sup>51</sup>

#### 4.6. Toponímia em língua portuguesa e o UNGEGN

As práticas de difusão de uma língua no mundo nem sempre são lineares. Do mesmo modo, a forma como estas são interpretadas é variável, seja pelo posicionamento político ou ideológico de quem recebe a língua, seja pela época em que se insere e pelo correspondente estado de coisas, em termos globais.

No que respeita à língua portuguesa, foram vários os papéis atribuídos à língua no seu trajeto de difusão, conforme é referido em Mateus (2014:32), desde “língua de prestígio” e “língua de expansão”, a partir do séc. XV, passando por “língua de subjugação cultural”, entre os séculos XVII e XIX e transformando-se numa “língua de unidade nacional”, com a emergência das colónias portuguesas enquanto países independentes. Este percurso trouxe-nos à atual “língua de tradição”, que nas palavras da autora é “um repositório de memórias que os povos que a falam reconhecem como parte do seu património, ao lado dos monumentos, das artes e ofícios, da música”. Bem como acontece com o restante léxico, também os

---

50 Machado (2003:1013): “Moneco, Do gr. *Mónoikos*, pelo lat. *Monoecus* (...). Como top.: *saxa Monóeci*, fortaleza de *Moneco*, na Ligúria, hoje Mónaco (...).”

51 «Cingalês», por especificar igualmente uma etnia que não é comum a todos os naturais do Seri Lanca e que, a somar, esteve recentemente envolvida em confrontos com a etnia tâmil, tem vindo a ser posto em causa, motivo pelo qual se admite já formas alternativas derivadas do topónimo atual e inclusivas para a generalidade dos naturais do Seri Lanca. O gentílico «seri-lanquês» é uma hipótese, se bem que ainda teórica.

topónimos são elementos “de tradição”, motivo pelo qual a sua difusão em língua portuguesa será, *a priori*, abraçada pelas comunidades, ainda que com resistências naturais, mais até ao nível local. A difusão da toponímia em português serve diretamente estas comunidades locais mas, simultaneamente, a lusofonia como um todo. Desta feita, é necessário um trabalho conjunto realizado pelos países que, mais do que concordarem com este propósito, nele se revejam.

Entre os países de língua oficial portuguesa é facilmente identificada uma lacuna no que respeita à existência de guias toponímicos articulados a submeter às Nações Unidas, especificamente ao UNGEGN,<sup>52</sup> grupo de trabalho afeto à padronização de topónimos (e nomes geográficos). Não terão sido ainda empregues esforços suficientes no sentido de colmatar a inexistência destes guias, podendo assumir-se uma importância secundarizada atribuída a este tipo de informação toponímica. Analisando os dados apresentados em Kerfoot (2006) e atualizando-os com recurso à informação disponível na página oficial do UNGEGN (United Nations Group of Experts on Geographical Names)<sup>53</sup>, verifica-se que esta posição, ou a ausência de uma estratégia bem definida, poderá não ser a mais adequada, uma vez que 41 países procederam já à apresentação de linhas orientadoras toponímicas próprias (ver Quadro A, em Anexos), dos quais 30 já com revisões e atualizações de primeiras versões.<sup>54</sup> No caso específico de Portugal, estamos perante um dos poucos “países mais desenvolvidos” – grupo formal da ONU que usa como unidade de medida a esperança média de vida à nascença –<sup>55</sup> sem linhas orientadoras toponímicas. No caso do Brasil, no seguimento de negociações com Angola e Moçambique, foram feitos alguns esforços para a criação da Divisão da Língua Portuguesa no Grupo de Peritos em Nomes Geográficos da ONU,<sup>56</sup> ainda que o processo não tenha alcançado o desenvolvimento inicialmente desejado: a “padronização dos nomes geográficos nacionais e internacionais”. Em Bustamante (2007) é feito um primeiro ensaio no sentido de agilizar este desiderato:

Durante o Curso Internacional de Toponímia das Nações Unidas, realizado em setembro de 2006 (em Maputo, Moçambique), representantes do Brasil, Moçambique e Angola iniciaram negociações e estão de acordo sobre a necessidade de apresentar esta proposta ao UNGEGN e também à conferência, a fim de fortalecer seus esforços individuais para promover a padronização dos nomes geográficos nacionais e internacionais e para atender aos interesses em comum da toponímia portuguesa mundial. (Bustamante 2007:2)

---

52 Traduzível para “Grupo de Peritos das Nações Unidas em Nomes Geográficos”, ou GPNUNG.

53 Informação disponível em <http://unstats.un.org/unsd/geoinfo/UNGEGN/wgGuidelines.html> (consultada em janeiro de 2016).

54 Destas quarenta e uma nações, no caso dos dados para a Rússia as linhas orientadoras foram apresentadas pela ainda URSS, em 1984 e 1987.

55 Tabela disponível na página das Nações Unidas, com o seguinte endereço: [http://www.un.org/esa/population/publications/UN\\_1982\\_Levels\\_and\\_Trends/UN1982\\_LevelsTrends\\_ch2\\_MDC.pdf](http://www.un.org/esa/population/publications/UN_1982_Levels_and_Trends/UN1982_LevelsTrends_ch2_MDC.pdf) (consultada em dezembro de 2015).

56 Artigo disponível em [http://www.ngb.ibge.gov.br/App\\_Doc/Proposal%20to%20Create%20the%20Portuguese%20Language%20Division\\_tradu%C3%A7%C3%A3o.pdf](http://www.ngb.ibge.gov.br/App_Doc/Proposal%20to%20Create%20the%20Portuguese%20Language%20Division_tradu%C3%A7%C3%A3o.pdf) (consultado em dezembro de 2015).





## **5. DADOS DO VOCABULÁRIO TOPONÍMICO**

Os dados do VT dividem-se em três grandes grupos. O dos topónimos correspondentes a lugares com relevância administrativa no espaço da CPLP; o dos nomes de países e de capitais de países; e o da informação relacional entre topónimos, por um lado, e entre topónimos e gentílicos, por outro. Adicionalmente, alguns topónimos relativos a referentes fora do espaço da CPLP estão incorporados no VT, por possuírem elevada frequência de uso entre os falantes de língua portuguesa e/ou por terem tradição lexicográfica forte, também com associação aos respetivos gentílicos atestados no VOC e com marcação interna da sua condição particular. Algumas unidades da toponímia menor no espaço da CPLP foram igualmente consideradas, tendo em vista o futuro alargamento da nomenclatura.

Relativamente aos gentílicos, importa sublinhar a ideia de que as referências para a sua atestação no léxico geral, mesmo que bastante abrangentes, não servem um propósito que se teria como ideal: a existência de pelo menos um gentílico atestado para cada topónimo. Tal atestação, plena, é ainda assim possível, pela aplicação de processos de sufixação previsíveis e estáveis.

O VT possui atualmente diversos casos de pares de topónimos homónimos, o que reforça a importância do recurso também enquanto ferramenta para a associação precisa de gentílicos a topónimos específicos, e vice-versa, bem como de cada topónimo a um nível hierárquico que o especifique como único. Note-se que os dados variam grandemente entre países e, dentro de cada país, entre níveis administrativos. Em alguns casos, aproximam-se da total unicidade (especificamente em Moçambique e, de uma forma geral, nos níveis administrativos maiores de cada país); noutros, existe considerável repetição (grande número de topónimos homónimos), sobretudo nos níveis administrativos mais baixos de Portugal, Brasil e Cabo Verde.

### **5.1. Países e capitais**

Os países do mundo – como as suas capitais – são unidades geográficas nomeáveis, de enorme visibilidade em qualquer língua viva, especialmente numa língua pluricêntrica, veículo de comunicação utilizado por um elevado número de falantes e possuidor de relevância destacada a nível global conferida por esse mesmo uso. É, assim, impreterível que uma língua como a portuguesa possua topónimos tratados de forma homogénea, desejavelmente únicos, que validem a nomeação consensual de todas as capitais e países, bem como o reconhecimento transversal de formas gentílicas associadas ao maior número possível destes topónimos, tendo por base referências históricas e, especificamente, da tradição lexicográfica, ainda que sem esquecer o uso.

Será também útil, de futuro, considerar topónimos que designem unidades políticas de tipo semelhante ao dos países mas que não integram a lista oficial de países da ONU por possuírem algum tipo de dependência política ou administrativa, mesmo não fazendo parte do território dos países que, de alguma forma, os controlam. Aqui se incluem províncias autónomas, de que são exemplo Alanda (presentemente

provincia da Finlândia) ou Curaçau (território autónomo dos Países Baixos), bem como Monserrate ou as Ilhas Caimão, territórios ultramarinos britânicos.

Fornecidas algumas referências para possível alargamento da base de dados toponímica que constitui o VT, importa focar os topónimos já tratados, disponibilizando em particular dados relativos a topónimos não regulares ou que possuem algum tipo de variação, com o devido enquadramento por tipo.

### 5.1.1. Países

Existem presentemente no mundo 194<sup>57</sup> países reconhecidos como tal pela ONU (Quadro B, em Anexos). Todos os topónimos correspondentes estão atestados no VT com marcação explícita de propriedades formais, grafados segundo a norma ortográfica vigente (AO90) e com marcação de variação nos casos em que esta ocorra. Entre os países considerados, existem 21 cujo nome ainda ou agora possui dupla grafia. Em dez destes, a variação decorre da altura da vogal tónica <e> ou <o> seguida de consoante nasal <m> ou <n> em posição de ataque (em início de sílaba), marcada na variedade brasileira com acento circunflexo e nas restantes variedades com acento agudo:

Quadro 5.1: Dupla grafia em nomes de países – casos com vogal tónica seguida de consoante nasal

#	CPLP, exceto Brasil	Brasil
1	Arménia	Armênia
2	Eslovénia	Eslovênia
3	Estónia	Estônia
4	Iémen	Iêmen
5	Letónia	Letônia
6	Macedónia	Macedônia
7	Mónaco	Mônaco
8	Polónia	Polônia
9	Quénia	Quênia
10	Roménia	Romênia

Nos restantes onze países em que por ora se admite dupla grafia, verificam-se fenómenos diferentes, que devem ser apontados. Um dos casos de variação decorre de questões extralinguísticas, sendo os dois topónimos em confronto válidos em todos os países da CPLP, um deles forma não adaptada:

Quadro 5.2: Dupla grafia em nomes de países – casos com motivação política e ideológica

#	CPLP
11	Myanmar/Birmânia

<sup>57</sup> Confronte-se os casos de Porto Rico e da Santa Sé.

Os remanescentes dez casos resultam das tradições lexicográficas distintas entre o Brasil e os restantes estados-membros da CPLP, explanada no próprio uso dos falantes, que rejeitam, no imediato, uma forma que não corresponda aos cânones que associam à sua variedade. Existem dois casos em que a terminação dos topónimos varia tendo por base a realização nasal final:

Quadro 5.3: Dupla grafia em nomes de países – casos com terminação nasal

#	CPLP, exceto Brasil	Brasil
12	Irão	Irã
13	Vietname	Vietnã

Existe um outro par de topónimos de radical morfológico comum mas com terminações diferentes, neste caso com distinção quanto ao número, já que no Brasil se atribui ao país o nome do príncipe Maurício de Nassau (singular) e nos restantes países da CPLP o topónimo perde a marca de singular própria do antropónimo que lhe dá origem. Ocorre uma extensão da marca de plural própria do número de ilhas que formam a nação (plural) para o topónimo (antes singular):

Quadro 5.4: Dupla grafia em nomes de países – caso decorrente da pluralização do topónimo

#	CPLP, exceto Brasil	Brasil
14	Maurícias	Maurício

Ocorre um caso em que a forma ortográfica adotada na variedade brasileira carece de adaptação, aproximando-se de uma simples transliteração da forma original:

Quadro 5.5: Dupla grafia em nomes de países – adaptação díspar

#	CPLP, exceto Brasil	Brasil
15	Bielorrússia	Belarus

Aos anteriores, acrescentam-se três casos de variação ortográfica decorrente daquilo que se entende como «produções orais distintas». Em cada um destes pares de topónimos, ocorre apenas um fone diferente em cada uma das duas sequências fonéticas, em tudo o resto iguais (estamos perante um «par mínimo»), motivando diferentes ortografias consoante a variedade em que são produzidas:

Quadro 5.6: Dupla grafia em nomes de países – variação decorrente de produção oral díspar

#	CPLP, exceto Brasil	Brasil
16	República Checa	República Tcheca
17	Turquemenistão	Turcomenistão
18	Usbequistão	Uzbequistão

Existem ainda dois casos de marcação de sílaba tónica diferente. Num deles, o topónimo é paroxítono na variedade brasileira e é proparoxítono nas restantes variedades (19); no outro caso, o topónimo é oxítono na variedade brasileira e paroxítono nas restantes variedades (20):

Quadro 5.7: Dupla grafia em nomes de países – diferente tonicidade

#	CPLP, exceto Brasil	Brasil
19	Domínica	Dominica
20	Madagáscar	Madagascar

Um caso de variação puramente ortográfica <s>/<c>, a ser eventualmente resolvido numa revisão futura, após nova avaliação feita pelos estados-membros da CPLP, está ainda atestado:

Quadro 5.8: Dupla grafia em nomes de países – variação puramente ortográfica

#	CPLP, exceto Brasil	Brasil
21	Singapura	Cingapura

### 5.1.2. Capitais

Relativamente às capitais de países, estão atestados 14 pares de topónimos com variação, sendo que o caso mais frequente (seis ocorrências) é aquele em que a sílaba tónica não é a mesma no Brasil e nas restantes variedades, existindo, no caso de Bagdade/Bagdá, efetiva supressão de uma sílaba no Brasil:

Quadro 5.9: Dupla grafia em nomes de capitais – diferente tonicidade

#	CPLP, exceto Brasil	Brasil
1	Lilongué	Lilongue
2	Malé	Male
3	Apia	Ápia
4	Moscovo	Moscou
5	Helsínquia	Helsinque
6	Bagdade	Bagdá

Tal como ocorre com os nomes de países, há dois casos em que a terminação dos nomes de capitais varia tendo por base a realização nasal final:

Quadro 5.10: Dupla grafia em nomes de capitais – casos com terminação nasal

#	CPLP, exceto Brasil	Brasil
7	Teerão	Teerã
8	Amesterdão	Amesterdã

Existem também dois casos em que a vogal final tem altura diferente na variante do Brasil, embora sejam exatamente opostas as realizações (9) e (10). *Iaundê* tem vogal final tónica fechada no Brasil, ao contrário do que ocorre nas restantes variedades, em que é aberta; ao passo que *Montevidéu* tem ditongo final tónico aberto no Brasil, enquanto a sua realização é fechada nas restantes variedades:

Quadro 5.11: Dupla grafia em nomes de capitais – casos com abertura de vogal final díspar

#	CPLP, exceto Brasil	Brasil
9	Iaundé	Iaundê
10	Montevideu	Montevidéu

Ocorrem três casos em que a terminação dos topónimos é diferente entre variedades, sendo que no caso de Tunes/Túnis, a vogal <i> em sílaba final átona torna necessária a acentuação gráfica da vogal precedente, tónica como nas restantes variedades, sendo a palavra sempre paroxítona:

Quadro 5.12: Dupla grafia em nomes de capitais – variação decorrente de produção oral díspar

#	CPLP, exceto Brasil	Brasil
11	Copenhaga	Copenhague
12	Vienciana	Vientiane
13	Tunes	Túnis

Ainda encontramos um caso de variação puramente ortográfica <s>/<c>, à semelhança do que acontece com o nome do país que lhes dá origem:

Quadro 5.13: Dupla grafia em nomes de capitais – variação puramente ortográfica

#	CPLP, exceto Brasil	Brasil
14	Cidade de Singapura	Cidade de Cingapura

### 5.1.2.1. Nomes de capitais não adaptados ortograficamente

Existem ainda capitais em cujo nome não existe variação mas cuja grafia se optou por manter fiel à língua de partida, que é, em dez dos casos, o inglês. Nos outros dois casos, a língua de partida é o francês, com a terminação *-ville* (*Brazzaville* e *Libreville*). A adaptação ao português seria problemática nesta altura, seja por ter de respeitar critérios de tradução pouco recomendáveis, como por exemplo a passagem de «-town», elemento não autónomo, para «cidade de»; ou a adaptação de topónimos derivados de antropónimos às regras de ortografia do português. Se bem que possíveis, estas adaptações não são prática comum ou consensual, logo, não são seguidas no VT. Assim, no VT conservaram-se 12 capitais enquanto formas não adaptadas, ou endónimos:

Quadro 5.14: Nomes de capitais de países não adaptados ortograficamente

#	Capital	Língua oficial	País
1	<i>Bridgetown</i>	Inglês	Barbados
2	<i>Brazzaville</i>	Francês	Congo
3	<i>Roseau</i>	Inglês	Domínica / Dominica
4	<i>Washington</i>	Inglês	Estados Unidos
5	<i>Libreville</i>	Francês	Gabão
6	<i>Georgetown</i>	Inglês	Guiana
7	<i>Kingston</i>	Inglês	Jamaica
8	<i>Wellington</i>	Inglês	Nova Zelândia
9	<i>Porto Moresby</i>	Inglês	Papua Nova Guiné
10	<i>Castries</i>	Inglês (e Francês)	Santa Lúcia
11	<i>Kingstown</i>	Inglês	São Vicente e Granadinas
12	<i>Freetown</i>	Inglês	Serra Leoa

### 5.1.3. Gentílicos associados a países

Os gentílicos fazem parte do léxico geral da língua, não estando diretamente atestados no VT. No entanto, existe uma relação derivacional entre os gentílicos e os topónimos que lhes dão origem, o que faz com que aqueles assumam, em alguns casos, a variação verificada nos topónimos-alvo. Desta feita, são disponibilizados na dissertação que se lê os casos em que existe variação na grafia de gentílicos de países, quando essa variação está associada ao uso numa determinada variedade da língua. Em sete dos nove casos de variação entre gentílicos de países, diferentemente grafados no Brasil e nos restantes membros da CPLP, as diferenças decorrem da derivação de topónimos – também eles – paralelos:

Quadro 5.15: Gentílicos decorrentes de topónimos paralelos

#	CPLP, exceto Brasil	Brasil	País
1	arménio	armênio	Arménia / Armênia
2	estónio	estônio	Estónia / Estônia
3	macedónio	macedônio	Macedónia / Macedônia
4	polaco	polonês	Polónia / Polônia
5	singapurense	cingapurense	Singapura / Cingapura
6	usbeque	uzbeque	Usbequistão / Uzbequistão

7	turquemenos	turcomenos	Turquemenistão / Turcomenistão
---	-------------	------------	--------------------------------

Relativamente aos gentílicos de países, existe também um caso atestado no VT em que a variação decorre do sufixo derivacional distinto empregue tendo por base o mesmo topónimo:

Quadro 5.16: Sufixo derivacional distinto empregue tendo por base o mesmo topónimo

#	CPLP, exceto Brasil	Brasil	País
8	canadiano	canadense	Canadá

Existe ainda um caso em que o gentílico utilizado no Brasil deriva de forma previsível do topónimo que lhe dá origem, contrariamente ao gentílico em uso para os restantes países da CPLP:

Quadro 5.17: Estado-unidense *versus* norte-americano

#	CPLP, exceto Brasil	Brasil	País
9	norte-americano	estado-unidense	Estados Unidos da América

Tal como ocorre com a variação excecional entre dois topónimos para um mesmo país (Myanmar/Birmânia), também os dois gentílicos possíveis daí derivados são distintos, correspondendo cada um deles a uma das duas designações toponímicas usadas para o referenciar. No caso do topónimo «Myanmar», o gentílico derivacional será ainda forma híbrida, por não estar o topónimo adaptado:

Quadro 5.18: Gentílicos derivados de topónimos distintos com um só referente geográfico (Myanmar/Birmânia)

#	CPLP	País
1	myanmarese/birmanês	Myanmar/Birmânia

Há mais cinco casos em que se admite mais do que um gentílico para um único país, independentemente da variedade, mas, aqui, o topónimo a que cada par de gentílicos se refere é exclusivo:

Quadro 5.19: Pares de gentílicos comuns a todas as variedades e derivados de um único topónimo

#	CPLP	País
1	conguês / conguense	Congo
2	conguês / conguense	República Democrática do Congo
3	guianense / guianês	Guiana
4	guineano / guineense	Guiné
5	equato-guineense / guineense	Guiné Equatorial

## 5.2. Divisões administrativas da CPLP (e dependências)

No que diz respeito ao tratamento de topónimos e verificação de gentílicos correspondentes no espaço da CPLP, fez-se o cruzamento de dados relativos a cada um dos países cuja toponímia se encontra atestada no VT (dados provenientes dos institutos nacionais de estatística próprios) com a informação que já existia no antigo Dicionário de Topónimos e Gentílicos, recurso do Portal da Língua Portuguesa que antecedeu o Vocabulário Toponímico. Os topónimos com relevância administrativa estão atestados, na sua totalidade, no VT, assim como alguns topónimos sem relevância administrativa: (i) quando se encontram disponíveis nos recursos já referidos; (ii) estão atestados noutras fontes de referência; (iii) ou têm uso frequente.<sup>58</sup> A verificação destas condições torna-os passíveis de receber tratamento homogéneo e coerente e assim passarem a fazer parte da nomenclatura do vocabulário. Os distritos do Brasil, por exemplo, estão incluídos já nesta versão do VT pela verificação destas condições, ainda que não possuam relevância administrativa.

### 5.2.1. Brasil

O VT inclui para o Brasil macrorregiões, estados, mesorregiões, microrregiões e municípios (mais os distritos, por agora). Atualmente, existem cinco macrorregiões no Brasil, que, embora num nível imediatamente abaixo do de «país», não possuem autonomia política. São elas: Norte, Nordeste, Centro Oeste, Sudeste e Sul. Cada uma das regiões é dividida em unidades federativas, que possuem, essas sim, autonomia. Vinte e seis estados e o Distrito Federal formam o conjunto das unidades federativas do Brasil, sendo que o Distrito Federal<sup>59</sup> possui apenas os três topónimos homónimos «Brasília», «Brasília» e «Brasília» nos níveis imediatamente abaixo (mesorregião, microrregião e município do Brasil).

A subdivisão «mesorregião» foi criada pelo IBGE para fins estatísticos e não possui relevância administrativa efetiva. Ainda assim, as 137 mesorregiões do Brasil são hierarquicamente relevantes na organização do país (contendo as microrregiões), motivo pelo qual foram consideradas válidas para inclusão no VT. As microrregiões, contendo os municípios (unidade mínima com relevância administrativa), são igualmente incorporadas.

Quadro 5.20: Subdivisões do Brasil (administrativas e não administrativas)

Subdivisão	Total de unidades
macrorregião	5
estado / unidade federativa	27
mesorregião	137
microrregião	558

58 A frequência é verificada em *corpora*, conforme se desenvolve em Almeida et al. (2013): “Cada *corpus* nacional terá no mínimo 30 milhões de palavras, distribuídas por sub-*corpora* com texto literário (20%), jornalístico (25%), legislativo e de sessões parlamentares (25%), técnico (saúde, educação, ambiente, pescas e agricultura: 25%) e de proveniência variada (5%).”

59 Unidade análoga aos estados, logo, atestada no VT ao nível destes, pese embora a sua regência esteja mais próxima da dos municípios do que de uma constituição estadual.



município	5570
distrito	~10283
bairro	n.d. (~14320)

### 5.2.2. Cabo Verde

Relativamente a Cabo Verde, as divisões relevantes são encabeçadas pelo nível «grupo» (Ilhas de Barlavento e Ilhas de Sotavento). Imediatamente abaixo, hierarquicamente, estão as nove «ilhas» do país, que se dividem em 22 concelhos. A unidade mínima com relevância administrativa é a freguesia.

Quadro 5.21: Subdivisões de Cabo Verde (administrativas e não administrativas)

Subdivisão	Total de unidades
grupo	2
ilha	9
concelho	22
freguesia	32
zona	n.d.
lugar	n.d.

### 5.2.3. Moçambique

No caso dos topónimos moçambicanos, deve ser feita a ressalva de que várias modificações e adaptações não decorreram da «conversão» de um topónimo tradicionalmente grafado num outro idioma nacional para a língua portuguesa mas, sim, da correção de grafias desatualizadas relativamente à norma ortográfica vigente para o português. «Chiôco», posto administrativo do distrito de Changara, província de Tete, é exemplo de um topónimo desatualizado adaptado às regras ortográficas vigentes.<sup>60</sup>

Quadro 5.22: Subdivisões de Moçambique (administrativas e não administrativas)

Subdivisão	Total de unidades
província	11
distrito	158
posto administrativo	405
localidade	~1141
bairro municipal / povoação	n.d.

<sup>60</sup> A vogal tónica <o> era marcada com um acento circunflexo em forma paroxítone, tornando o topónimo agramatical segundo as regras do AO90 (paroxítone com vogal <o> tónica acentuada graficamente, não se tratando de <pôde>, forma excecional, ou de <fôrma>, forma facultativa mas, ainda assim, também excecional.

#### 5.2.4. Portugal

Para Portugal, foram considerados os dados toponímicos ao nível das freguesias, municípios e distritos do país, bem como os níveis NUTS.<sup>61</sup> Uma primeira fase do trabalho consistiu na verificação dos topónimos que constituem a nomenclatura dos CTT, informação que possibilitou uma mais sólida análise dos topónimos registados nos Censos 2011<sup>62</sup> – disponibilizados pelo INE –, sendo feita correção de erros e lacunas existentes e atualizando os casos, específicos, em que tal foi necessário, em conformidade com as regras do AO90. Numa segunda fase, esses topónimos foram confrontados com os dados já existentes em versão do VOP anterior à atual.<sup>63</sup> Destes cruzamentos de dados, verificações de erros e inconsistências, correções de topónimos mal-formados ou desatualizados, incorporação de níveis NUTS e homogeneização de entradas resultou o conjunto atual de topónimos que para Portugal integram o VT.

O VT atesta para Portugal 3091 freguesias, 308 municípios, 29 distritos/ilhas (25 unidades NUTS III), sete unidades NUTS II e três unidades NUTS I, estando estes dados gerais, exceto os referentes à NUTS, quantitativamente, em conformidade com os disponibilizados na CAOP – Carta Administrativa Oficial de Portugal em vigor, de 2015.<sup>64</sup>

Quadro 5.23: Subdivisões de Portugal (administrativas e não administrativas)

Subdivisão	Total de unidades
NUTS I	3
NUTS II	7
NUTS III / distrito	NUTS III (25) / distritos (29)
município	308
freguesia	3091
lugar	n.d.
local	n.d.

61 Para a atual União Europeia, de que Portugal faz parte, foi criada a NUTS pelo “Regulamento (CE) n.º 1059/2003 do Parlamento Europeu e do Conselho de 26 de Maio de 2003 relativo à instituição de uma Nomenclatura Comum das Unidades Territoriais Estatísticas (NUTS)”, com atualizações até à última versão: o “Regulamento (UE) n.º 868/2014 da Comissão de 8 de agosto de 2014, que altera os anexos do Regulamento (CE) n.º 1059/2003 do Parlamento Europeu e do Conselho relativo à instituição de uma Nomenclatura Comum das Unidades Territoriais Estatísticas (NUTS)”. A NUTS considera três níveis hierárquicos: NUTS I, NUTS II e NUTS III. Estes níveis de identificação territorial corresponderam inicialmente a regiões meramente estatísticas mas acabaram por ganhar relevância administrativa, desde logo pela atribuição de verbas, a nível europeu, através de estruturas neles baseadas. Por este motivo, os níveis NUTS são incluídos no VT, ainda que o nível III seja, de alguma forma, problemático, pela sobreposição total ou parcial de áreas geográficas NUTS III com distritos. No momento, «distrito» e «NUTS III» são inseridos no VT ao mesmo nível, sendo concorrentes.

62 Veja-se [http://censos.ine.pt/xportal/xmain?xpgid=censos2011\\_apresentacao&xpid=CENSOS](http://censos.ine.pt/xportal/xmain?xpgid=censos2011_apresentacao&xpid=CENSOS) (página consultada em fevereiro de 2016).

63 O VOP tem como fonte principal o Vocabulário da Língua Portuguesa, de Francisco Rebelo Gonçalves (1966).

64 Carta disponível em [http://www.dgterritorio.pt/static/repository/2015-07/2015-07-24104933\\_b511271f-54fe-4d21-9657-24580e9b7023\\$\\$D95E63D4-0C6F-47AD-9A9D-2396C1735714\\$\\$201AB9C9-EDF7-492C-8334-4D8998D68EA2\\$\\$file\\$\\$pt\\$\\$1.pdf](http://www.dgterritorio.pt/static/repository/2015-07/2015-07-24104933_b511271f-54fe-4d21-9657-24580e9b7023$$D95E63D4-0C6F-47AD-9A9D-2396C1735714$$201AB9C9-EDF7-492C-8334-4D8998D68EA2$$file$$pt$$1.pdf) (página consultada em fevereiro de 2016).

### 5.2.5. São Tomé e Príncipe

Em São Tomé e Príncipe, as divisões administrativas são feitas partindo de duas províncias (ou ilhas), sete distritos<sup>65</sup> e 480 estabelecimentos. Estas são as unidades do país que possuem relevância administrativa, estando neste momento tratados e integrados no VT todos os topónimos que delas fazem parte.

Quadro 5.24: Subdivisões administrativas de São Tomé e Príncipe

Subdivisão	Total de unidades
províncias/ilhas	2
distritos	7
estabelecimentos	480

### 5.2.6. Timor-Leste

Timor-Leste é dividido administrativamente em 13 distritos, 67 subdistritos e 442 sucos. Neste caso particular, as dificuldades de adaptação de todos os topónimos com relevância administrativa às regras ortográficas da língua portuguesa foram mais sentidas, pelo maior afastamento entre as especificidades do sistema gráfico do tétum e aquelas do português. Não obstante, tendo por base critérios alcançados após discussão com as autoridades e com especialistas do país, chegou-se à primeira versão de um *corpus* toponímico leste-timorense incluindo todos os topónimos que possuem relevância administrativa grafados em língua portuguesa. Estes dados foram então criteriosamente trabalhados e integrados no VT.

Quadro 5.25: Subdivisões administrativas de Timor-Leste

Subdivisão	Total de unidades
distritos	13
subdistritos	67
sucos	442

### 5.2.7. Angola e Guiné-Bissau

Relativamente a Angola e à Guiné-Bissau, não foi ainda desenvolvida para publicação a toponímia em língua portuguesa seguindo os critérios para os restantes países da CPLP. Esses dados foram já, em parte, trabalhados internamente, tendo em vista a sua apresentação aos países interessados para posterior discussão, convenientemente fundamentada, com as autoridades competentes e os especialistas de cada país mandatados para trabalhos nesse sentido.

---

65 Seis distritos da província de São Tomé (Água Grande, Mé Zóchi, Cantagalo, Caué, Lembá, Lobata) e um (Pagué) da Ilha do Príncipe, região autónoma.

Quadro 5.26: Subdivisões administrativas de Angola

Subdivisão	Total de unidades
província	n.d.
município	n.d.
comuna	n.d.

Quadro 5.27: Subdivisões administrativas da Guiné Bissau

Subdivisão	Total de unidades
região	n.d.
setor	n.d.

### 5.3. Uniformização e relações entre unidades toponímicas

Todos os topónimos com relevância administrativa no espaço da CPLP, atestados no VT, bem como aqueles que não a possuem mas que nele estejam já integrados, passam, portanto, a estar uniformizados ortograficamente segundo critérios que respeitam a norma ortográfica vigente nas áreas políticas que os contêm. Quer isto dizer que todos são verificados e, quando necessário, adaptados a uma norma ortográfica única e comum. Tanto no caso da toponímia maior para o espaço da CPLP, como no caso dos topónimos que identificam países e suas capitais, uma estrutura hierárquica e homogênea é conseguida através da atribuição de um código único que permite identificar inequivocamente qualquer referente geográfico atestado, bem como estabelecer relações lógicas entre topónimos e gentílicos (sejam estes derivacionais ou não). Desta forma, topónimos – e gentílicos – potencialmente homófonos são facilmente distinguidos e, mesmo nos casos em que sejam homónimos, passam a ter enquadramento numa posição hierárquica exclusiva, o que lhes confere um espaço único na organização do VT.

### 5.4. Codificação

A relação entre as unidades toponímicas é então possível, ao nível dos topónimos, tendo por base um código alfanumérico atribuído a cada topónimo atestado e, entre topónimos e gentílicos, estabelecendo uma ligação, por meio de informação interna na matriz de cada constituinte do par topónimo-gentílico, dando conta – na entrada correspondente ao gentílico – da ocorrência da forma de partida (topónimo) e – na entrada do topónimo – da derivação que deste resulta ou da forma não derivacional que lhe está associada (gentílico). Cada código é único, de visualização apenas interna, e permite identificar inequivocamente o referente geográfico nomeado, bem como as relações que este tem com outros constituintes. No caso dos códigos para os nomes dos países, é usado o ISO 3166-1,<sup>66</sup> com a recomendada codificação «Alpha-2 code», ou seja, com uma codificação única constituída por duas letras (de notar que são igualmente possíveis codificações

66 International Standard for Country Names, disponível em [http://www.iso.org/iso/country\\_codes](http://www.iso.org/iso/country_codes).

«Alpha-3 code» ou «numéricas».<sup>67</sup> Ao passo que um código «Alpha-2» permite identificar um topónimo, um código composto que integre o «Alpha-2» e códigos para outros níveis permite identificar igualmente um topónimo único mas também a sua localização numa estrutura hierárquica. Por exemplo, para o topónimo «Japão», teremos o código «Alpha-2» JP, enquanto parte de um código alfanumérico único maior: 142 (continente\_Ásia) 030 (subcontinente\_Ásia Oriental) JP (país\_Japão) – 142 030 JP. No caso do nome das capitais dos países, não existe uma codificação-padrão utilizada transversalmente em termos internacionais, tendo sido criada para o VT uma primeira versão de uma codificação alfabética composta por três letras, normalmente correspondentes aos três primeiros caracteres do nome de cada capital (em língua portuguesa). No caso de os três caracteres iniciais de duas ou mais capitais serem os mesmos, o quarto carácter substitui o terceiro na codificação da capital do país que alfabeticamente for posterior (e.g. *BER* (Berlim); *BEN* (Berna). Se, ainda assim, o código formado for igual a outro, é selecionada a letra seguinte (e.g. *BAN* (Banjul); *BAG* (Bagdade); *BAU* (Bangui); *BAE* (Banguécoque). No caso das capitais com nome composto, a codificação assume a primeira letra do primeiro constituinte e as duas primeiras do último. Se, mesmo assim, resultarem dois ou mais códigos iguais, a terceira letra do código passa a ser a terceira letra do último constituinte na codificação da capital do país que seja alfabeticamente posterior (e.g. *SJO* (São João); *SJS* (São José).

Assim, sendo o nível menor na classificação de topónimos não pertencentes à CPLP o de «capital de país», o código mais extenso possível para este subconjunto terá onze caracteres. Como exemplo, a capital japonesa passa a ser identificável através do código único 140 030 JP TOQ.

A codificação de todos os topónimos incluídos no VT respeita então uma estrutura comum que inclui (i) uma identificação única, ou *ID* (por exemplo, de 1 a 70.000); (ii) a forma ortográfica do topónimo, disponível para visualização por qualquer utilizador; (iii) o código do topónimo, que permite localizá-lo e relacioná-lo com outros topónimos, numa estrutura que é hierárquica; (iv) o nível do topónimo, de uma tabela com níveis comuns a todos os topónimos (por exemplo, abaixo de «país» está o nível único «6»); (v) o nome efetivo associado ao nível a que corresponde um topónimo (por exemplo, «distrito»); e (vi) um valor para a (im)possibilidade de ocorrência de artigo com o topónimo.<sup>68</sup>

---

67 Existe também a codificação 3166-2, para identificação das subdivisões dos países (províncias ou estados), nível que não é tido em conta nesta versão do VT. Adicionalmente, a codificação 3166-3, que diz respeito a nomes de países outrora classificados segundo o ISO 3166-1, é excluída deste recurso, por motivos vários, que são explicados em detalhe em [http://www.iso.org/iso/country\\_codes\\_glossary.html](http://www.iso.org/iso/country_codes_glossary.html) (consulta feita em março de 2016).

68 O artigo é tido como critério de identidade lexical, logo, é aqui incluído.



## 6. CONCLUSÕES E TRABALHO FUTURO

A publicação do Vocabulário Toponímico (VT), segundo se pôde aferir no decorrer da redação desta dissertação, significa mais do que a mera disponibilização de uma lista de palavras, *per se*. Ela contribui para o esclarecimento de questões toponímicas recorrentes aos mais diversos níveis, esclarecimento esse que se torna possível através da consolidação normativa, pela prescrição de topónimos idealmente únicos (sem variação meramente ortográfica) que têm linha de conta uma ortografia comum sem que seja esquecida a singularidade de cada variedade do português e as ligações identitárias entre cada população e os topónimos que lhe são caros. Conclui-se que a existência do VT, seguindo o que ocorre com os dados do próprio VOC, releva as variedades até aqui sub-representadas na lexicografia de língua portuguesa, com a intenção de as colocar a par das variedades brasileira e portuguesa, de tradição lexicográfica mais forte também ao nível da toponímia.

### 6.1. Abrangência e utilidade do recurso

Tal como foi explicitado no decorrer desta dissertação, o VT contempla necessariamente todos os topónimos com relevância administrativa no espaço da CPLP, bem como aqueles que designam países (reconhecidos pela ONU) e suas capitais. Não exclui necessariamente todos os outros, ou seja, inclui aqueles que, pela relevância que possuam (maioritariamente de tradição ou de uso), sejam uma mais-valia para o utilizador da língua. Além das propriedades formais de cada topónimo, os dados são complementados por informação gentilícia e por relações entre todas as unidades atestadas. Não existe uma entrada no vocabulário que não possua uma ligação a outra. Na realidade, o VT é uma rede de unidades toponímicas, finita e estruturalmente consolidada mas com larga margem para desenvolvimentos futuros. O alcance de um recurso com estas características torna-se evidente, na medida em que os dados, abertos, servem de azimute no desenvolvimento de estudos e recursos toponímicos em sociologia, linguística, geografia ou história, apenas para nomear algumas áreas do conhecimento humano que do VT podem retirar mais-valias. Na base desta possibilidade está o facto de o VT permitir adequadas referências para planeamento territorial, com difusão uniforme de todos os dados ao nível toponímico que fazem parte do espectro do trabalho: novamente, aqueles que presentemente assumem maior relevância em cada um dos países da CPLP.

O VT, ainda que não seja uma ferramenta topográfica, torna a adequada georreferenciação em língua portuguesa uma realidade – pela sistematização cartográfica que possibilita –, dirimindo ambiguidades decorrentes da coexistência de sistemas de referenciação com diferentes metodologias e com atestações díspares, com elevado grau de opacidade entre si. Daqui decorrem mais-valias também em termos económicos, pela diminuição de trabalho duplicado por parte de diferentes instituições (não interligadas) que hoje gastam, necessariamente, tempo na pesquisa e cruzamento de referências para publicação de dados toponímicos que sejam minimamente homogêneos, transversalmente reconhecíveis.

Por todo o trabalho desenvolvido na elaboração do VT e partilhado nesta dissertação, com exposição das questões mais dificultosas e da metodologia previamente desenhada, bem como da evolução do recurso com auxílio de linhas teóricas bem fundamentadas para a adequada persecução de um objeto toponímico útil à comunidade internacional e gratuito para todos quantos tenham interesse por questões relativas à toponímia, torna-se estimulante iniciar exercícios que permitam esboçar futuros projetos decorrentes – total ou parcialmente – dos dados disponibilizados no Vocabulário Toponímico.

## 6.2. Desenvolvimentos e aplicações futuros

O VT é também um instrumento de política linguística. Se preferirmos, é o resultado de decisões veiculadas politicamente, cujo objetivo passa pelo reforço da política de difusão da língua portuguesa enquanto património cultural comum. Neste sentido, o VT assume por agora o seu papel de veículo normativo para a toponímia em língua portuguesa no espaço da CPLP mas, previsivelmente, alargará as suas competências e afirmar-se-á como recurso de âmbito mundial, ainda que com identidade equiparada à de outros recursos extra-CPLP. Tal como existem hoje inúmeras ações de promoção e divulgação da língua portuguesa por países em que o português não é língua oficial (nomeadamente realizadas pelo Camões – Instituto da Cooperação e da Língua, de Portugal, ou pela Rede Brasil Cultural, do Brasil), também o alargamento do VT com dados relativos à toponímia internacional (extra-CPLP) desempenhará importante papel de difusão do idioma. A hipótese de que o VT serve de ponto de partida para o alargamento desta promoção é verificada pela abrangência e sistematização dos dados já trabalhados na plataforma VOC.

Com os dados gerais disponibilizados pelo VT é possível criar ferramentas computacionais, nomeadamente de modelização, para estudo e desenvolvimento de trabalhos que estejam, de alguma forma, relacionados com topónimos. Uma das possíveis aplicações dos dados será uma modelização que permita formar um gentílico derivado para cada topónimo que ainda não o possua, para isso recorrendo a uma base morfológica e a um sufixo produtivo, de que são exemplo *-ense* ou *-ês*, e tendo em conta a terminação da base, o índice temático do topónimo e a sua tonicidade. Esta modelização é suficientemente abrangente para ser produtivamente alargada à toponímia menor, depois de estar esta convenientemente tratada, bem como a topónimos fora do espaço da CPLP que venham a ser adaptados à língua portuguesa e integrados no VT ou noutra recurso que, partindo deste, se desenvolva.

Os grupos de trabalho da ONU desenvolvem trabalho aturado para a toponímia, com avaliação das especificidades de cada país, sendo que cada nação contribui com um manual próprio elaborado por especialistas na área mas seguindo, de forma mais ou menos rigorosa, diretivas emanadas em concertação. Os países da CPLP, neste sentido, têm ainda de desenvolver esforços efetivos, recorrendo a dados próprios estruturados, comuns, com metodologia de organização consolidada. O VT serve de respaldo para a apresentação de linhas orientadoras e normas aos grupos de trabalho, uma vez que os dados do recurso, qualitativamente sólidos e quantitativamente relevantes, servem de base para a implementação de diretivas toponímicas a nível internacional.



De futuro, e de forma faseada, trabalhando em colaboração com os organismos competentes em cada um dos países cuja toponímia venha a ser tratada em língua portuguesa, será possível alargar de forma considerável os dados toponímicos hoje nomenclatura do VT. Poder-se-á aplicar a mesma metodologia adotada no VT para os países da CPLP, começando por tratar os topónimos de nível hierarquicamente superior com relevância administrativa em cada um dos países considerados. Dependendo do país a tratar, da língua oficial e do alfabeto nacional, bem como de outras especificidades linguísticas e políticas, os processos aplicados por forma a constituir uma base de dados toponímica em língua portuguesa serão necessariamente diferentes. Ainda assim, tem-se como perfeitamente concretizável o alargamento dos dados do VT, ou a criação de um outro recurso, com a toponímia específica de outros países.

Será relevante apostar no desenvolvimento do VT pela integração igualmente faseada de dados de toponímia menor mais alargados. Este tipo de trabalho trará novos constrangimentos e dificuldades metodológicas mas permitirá uma ampliação considerável dos dados. Se considerarmos que, por exemplo para Portugal, existem neste momento cerca de 3500 topónimos atestados no VT até ao nível «freguesia», com o acréscimo decorrente da padronização de nomes geográficos ao nível «lugar», este número mais do que triplica.



## BIBLIOGRAFIA

- Almeida, Gladis Maria de Barcellos, José Pedro Ferreira et al. (2013), “Vocabulário Ortográfico Comum (VOC): constituição de uma base lexical para a língua portuguesa”, São Paulo, in: Estudos Linguísticos 42 (1).
- Ávila, Patrícia (2008), *A Literacia dos Adultos: Competências-Chave na Sociedade do Conhecimento*, Lisboa, Celta.
- Bosque, Ignacio, Violeta Demonte (1999), *Gramática descriptiva de la lengua española*, Madrid, Colección Nebrija y Bello.
- Bourdieu, Pierre (1998[1982]), *O que Falar quer Dizer. A Economia das Trocas Linguísticas*, Lisboa, Difel.
- Bright, William (2003), “What IS a Name? Reflections on Onomastics”, Taipé, Institute of Linguistics, Academia Sinica.
- Brito, Regina Helena, Moisés de Lemos Martins (2004), “Considerações em torno da relação entre língua e pertença identitária em contexto lusófono”, Anuário Internacional de Comunicação Lusófona, n.º 2, São Paulo, Federação Lusófona de Ciências da Comunicação.
- Bustamante, Ana Maria Goulart, Moema José de Carvalho Augusto (2007), “Proposta para a criação da Divisão da Língua Portuguesa no Grupo de Peritos em Nomes Geográficos da ONU”, Nova Iorque, IBGE.
- Carvalhinhos, Patrícia de Jesus (2002), “Onomástica e Lexicologia: O Léxico Toponímico como Catalisador de Fundo de Memória. Estudo de caso: os sociotopónimos de Aveiro (Portugal)”, São Paulo, Revista USP.
- Carvalhinhos, Patrícia de Jesus (2008), “Estudos de Onomástica em língua portuguesa no Brasil: perspectivas para inserção mundial”, São Paulo, Simpósio Mundial de Estudos de Língua Portuguesa.
- Casteleiro, João Malaca (Coord.) (2001), *Dicionário da Língua Portuguesa Contemporânea*, Lisboa, Academia das Ciências de Lisboa e Editorial Verbo.
- Castells, Manuel (1999), *O poder da identidade*, Volume II, Tradução de Klaus Brandini Gerhardt, São Paulo, Paz e Terra.
- Castells, Manuel (2003), “Paraísos comunais: identidade e significado na sociedade em rede”, Rio de Janeiro, DP&A.
- Castells, Manuel, Gustavo Cardoso (Orgs.) (2005), *A Sociedade em Rede, Do Conhecimento à Acção Política*, Lisboa, Imprensa Nacional – Casa da Moeda.
- Correia, Margarita, José Pedro Ferreira (2013), “O papel dos dicionários e dos vocabulários ortográficos na constituição da norma”, em *O português no século XXI: cenário geopolítico e sociolinguístico*, São Paulo, Parábola.
- Correia, Margarita (2015), “Topónimos e ortografia – os topónimos no espaço público”, Benguela, Congresso sobre as Línguas Nacionais de Angola, Piaget Benguela.
- Costa, Margarida Faria da, Sofia Pereira da Silva (Coords.) (2004), *Grande Dicionário da Língua Portuguesa*, Porto, Porto Editora.
- Cunha, Celso, Lindley Cintra (1984): “Nova Gramática do Português Contemporâneo”, 17.ª Edição (2002), Lisboa, Edições João Sá da Costa
- Darrigol, Adeline (2015), *Política Linguística e Toponímia na Guiné Espanhola*, Bruxelas, 11.ª Jornada Científica REALITER, Academia Real da Bélgica. Tradução: Jales Josino da Rocha Filho, Boletim *a folha*, n.º 49 – outono de 2015, Comissão Europeia.
- Dick, Maria Vicentina de Paula do Amaral (1990), “A motivação toponímica e a realidade brasileira”, São Paulo, Geo Arquivo do Estado São Paulo.
- Dick, Maria Vicentina de Paula do Amaral (1990), “Toponímia e Antroponímia no Brasil”, São Paulo, Coletânea de

- Estudos. 2.<sup>a</sup> edição.
- Encarnação, José d' (2010), "A Toponímia, Fonte Histórica também para a República", em *Revista Biblios*, Volume VII – (2.<sup>a</sup> Série) 73-82, Coimbra, Faculdade de Letras, Universidade de Coimbra.
- Ferreira, Aurélio Buarque de Holanda (1999), *Novo Aurélio Século XXI: o dicionário da língua portuguesa* – 3.ed., Rio de Janeiro, Nova Fronteira.
- Ferreira, José Pedro, Maarten Janssen et al. (2012), "The Common Orthographic Vocabulary of the Portuguese Language: a set of open lexical resources for a pluricentric language", Istambul, LREC.
- Ferreira, José Pedro, Margarita Correia (2014), "Vocabulário Ortográfico do Português: descrição do projeto", in: Garcés Gómez, María Pilar (ed.), "Lexicografía teórica y aplicada" (Anexo de Revista de Lexicografía), Corunha, Universidade da Coruña.
- Ferreira, Carlos do Nascimento (2013), *Sendin – Tierra de Miranda: Geografia e Toponímia*, Lisboa, Âncora Editora.
- Garnelo, Fernando Bello (2001), "La toponímia de la zona arqueológica de las Médulas (León)", Universidad de León
- Gonçalves, Francisco Rebelo (1966), *Vocabulário da Língua Portuguesa*, Coimbra, Coimbra Editora.
- Handloff, Robert Earl (ed.) (1991), *Cote d'Ivoire: a country study*, Federal Research Division, Washington, Library of Congress.
- Houaiss, Antônio, Mauro de Salles Vilar (2001), *Dicionário Houaiss da Língua Portuguesa*, Instituto Antônio Houaiss de Lexicografia e Banco de Dados da Língua Portuguesa S/C Ltda., Rio de Janeiro, Objetiva.
- Houtman, Gustaaf (1999), "Mental Culture in Burmese Crisis Politics – Aung San Suu Kyi and the National League for Democracy", Tóquio, Series no. 33, ILCAA.
- Kerfoot, Helen, Eeva Maria Närhi (2006), "Manual for the national standardization of geographical names", Nova Iorque, United Nations Publication.
- Landry, Rodrigue, Richard Bourhis (1997), "Linguistic Landscape and Ethnolinguistic Vitality: An Empirical Study", *Journal of Language and Social Psychology*, Vol. 16, No. 1, Sage Publications, Inc.
- Machado, José Pedro (2003), *Dicionário Onomástico Etimológico da Língua Portuguesa*, Lisboa, Livros Horizonte.
- Mateus, Maria Helena Mira (2014), *A Língua Portuguesa: Teoria, Aplicação e Investigação*, Lisboa, Edições Colibri.
- Mattoso, José (1998), *A identidade nacional*, Gradiva, Lisboa.
- Mello, Antônio Augusto Souza (2000), "Estudo Histórico da Família Lingüística Tupi-Guarani – Aspectos Fonológicos e Lexicais", Florianópolis.
- Oriol, Michel (1985), "Appartenance linguistique, destin collectif, décision individuelle", in: *Cahiers Internationaux de Sociologie*, Paris, PUF. V. LXXXIX.
- Ormeling, Ferjan (1993), "Exonyms in Cartography", UNGEGN Training Course in Toponymy for Southern Africa, Pretória, Department of Geography, University of Pretoria.
- Quivy, Raymond e Luc Van Campenhoudt (1992), *Manual de Investigação em Ciências Sociais*, Lisboa, Gradiva.
- Rocha, José António Oliveira, Gonçalves Zavale (2015), "O Desenvolvimento do Poder Local em África: O caso dos municípios em Moçambique", *Caderno de Estudos Africanos*, Lisboa, Centro de Estudos Internacionais.
- Rodrigues, Patrícia Alexandra Pascoal (2008), "Dados e Práticas Festivas no(s) Espaço(s) Público(s) Lisboaeta(s) (1974-2005)", Lisboa, VI Congresso Português de Sociologia.
- Santos, Cláudio João Barreto dos (2008), *Geonímia do Brasil: A padronização dos nomes geográficos num estudo de caso dos municípios fluminenses*, Universidade Federal do Rio de Janeiro, Rio de Janeiro, IGEO.
- Silva, José Manuel Azevedo e (2003), "O modelo pombalino de colonização da Amazônia", Coimbra, Revista de

História e Sociedade da Cultura.

Tichelaar, Tjeerd (2002), “Toponymy and language”, Frankfurt, United Nations Publication.

UNGEGN Working Group on Definitions (1972), “A glossary of technical terminology for employment in the standardization of geographical names”, Second United Nations Conference on the Standardization of Geographical Names, Londres.

Vasconcellos, José Leite de (1966), *Lições de Filologia Portuguesa*, Quarta Edição, Rio de Janeiro, Livros de Portugal.

Woodman, Paul (2009), “The Nature of the Endonym”, UNGEGN Working Group on Exonyms, Working Paper No. 1, Twenty-fifth session, Nairóbi.

## WEBGRAFIA

Acreano ou acriano, petição pública: <http://www.peticaopublica.com.br/pview.aspx?pi=BR87797>

Artigo sobre renomeação de ruas tendo por base uma motivação política e/ou ideológica: <https://www.theguardian.com/cities/2016/jun/28/streets-new-names-airbrush-politics-renaming-roads>

Carta Administrativa Oficial de Portugal (CAOP): [http://www.dgterritorio.pt/static/repository/2015-07/2015-07-24104933\\_b511271f-54fe-4d21-9657-24580e9b7023\\$\\$D95E63D4-0C6F-47AD-9A9D-2396C1735714\\$\\$201AB9C9-EDF7-492C-8334-4D8998D68EA2\\$\\$file\\$\\$pt\\$\\$1.pdf](http://www.dgterritorio.pt/static/repository/2015-07/2015-07-24104933_b511271f-54fe-4d21-9657-24580e9b7023$$D95E63D4-0C6F-47AD-9A9D-2396C1735714$$201AB9C9-EDF7-492C-8334-4D8998D68EA2$$file$$pt$$1.pdf)

Censos 2010 Brasil: <http://censo2010.ibge.gov.br/>

Censos 2010 Cabo Verde: <http://www.ine.cv/censo/censo2010.aspx>

Censos 2007 Moçambique: <http://www.ine.gov.mz/operacoes-estatisticas/censos/censo-2007>

Censos 2011 Portugal: [http://censos.ine.pt/xportal/xmain?xpid=CENSOS&xpgid=censos2011\\_apresentacao](http://censos.ine.pt/xportal/xmain?xpid=CENSOS&xpgid=censos2011_apresentacao)

Censos 2012 São Tomé e Príncipe: <http://www.ine.st/2012.html>

Censos 2015 Timor-Leste: <http://www.statistics.gov.tl/pt/>

Códigos de países (ISO): [http://www.iso.org/iso/country\\_codes](http://www.iso.org/iso/country_codes)

Dados CIA World Factbook para Timor-Leste: <https://www.cia.gov/library/publications/the-world-factbook/geos/tt.html>

Competências de juntas de freguesia: <http://www.cm-lisboa.pt/servicos/por-temas/reforma-administrativa-de-lisboa/competencias-das-juntas-de-freguesia>

Dados históricos de utilização de artigo em «Guiné»: <http://www.corpusdoportugues.org/>

Deliberação n.º 6/AML/2014, de 22 de janeiro de 2014: [http://www.cm-lisboa.pt/fileadmin/DOCS/Reforma\\_Administrativa\\_Lisboa/Compet%C3%A2ncias/Deliberacao\\_Proposta\\_final\\_4\\_2014.pdf](http://www.cm-lisboa.pt/fileadmin/DOCS/Reforma_Administrativa_Lisboa/Compet%C3%A2ncias/Deliberacao_Proposta_final_4_2014.pdf)

Divisão da Língua Portuguesa para o UNGEGN: [http://www.ngb.ibge.gov.br/App\\_Doc/Proposal%20to%20Create%20the%20Portuguese%20Language%20Division\\_tradu%C3%A7%C3%A3o.pdf](http://www.ngb.ibge.gov.br/App_Doc/Proposal%20to%20Create%20the%20Portuguese%20Language%20Division_tradu%C3%A7%C3%A3o.pdf)

Glossário para códigos de países (ISO): [http://www.iso.org/iso/country\\_codes\\_glossary.html](http://www.iso.org/iso/country_codes_glossary.html)

INE, divisões administrativas, incluindo definição de «lugar»: [https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine\\_cont\\_inst&INST=6251013&xlang=pt](https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_cont_inst&INST=6251013&xlang=pt)

Linhas orientadoras UNGEGN: <http://unstats.un.org/unsd/geoinfo/UNGEGN/wgGuidelines.html>

Organização Internacional de Normalização (ISO): <http://www.iso.org>

Países considerados pela ONU como sendo os “mais desenvolvidos” e critérios para esta definição:

[http://www.un.org/esa/population/publications/UN\\_1982\\_Levels\\_and\\_Trends/UN1982\\_LevelsTrends\\_ch2\\_MDC.pdf](http://www.un.org/esa/population/publications/UN_1982_Levels_and_Trends/UN1982_LevelsTrends_ch2_MDC.pdf)

Reformas e acordos ortográficos do português: <http://www.portaldalinguaportuguesa.org/?action=acordo-historia>



Topónimo exclusivo «Timor-Leste»: [https://en.wikipedia.org/wiki/East\\_Timor#Etymology](https://en.wikipedia.org/wiki/East_Timor#Etymology)

## ANEXOS


































Quadro A: Países com linhas orientadoras toponímicas para uso internacional submetidas à UNGEGN

<b>País</b>	<b>Data da primeira submissão</b>	<b>Data da última submissão</b>
África do Sul	1984	1992
Alemanha	1981	2011
Algéria	2000	data única
Austrália	1992	2002
Áustria	2006	2012
Bélgica	2009	data única
Canadá	1981	2000
Chile	2007	data única
China	1982	1992
Chipre	2002	data única
Coreia do Sul	2002	2012
Croácia	2007	data única
Dinamarca	2007	2012
Eslováquia	1996	2000
Eslovênia	1995	data única
Espanha	1992	2006
Estados Unidos	1982	1989
Estônia	1992	2012
Finlândia	1981	2012
França	1981	2007
Grécia	1981	2000
Hungria	1981	2012
Indonésia	2007	data única
Irão	2000	2009
Irlanda	1992	1994
Islândia	1982	data única
Itália	1987	2004
Japão	1987	2007
Noruega	1982	2007
Países Baixos	1986	2012
Polónia	1996	2011
Reino Unido	1982	2009
República Checa	1994	2007
Roménia	1991	1994
Suécia	1981	2007
Suriname	1981	1985
Suíça	1982	data única
Tailândia	2002	data única
Turquia	1982	data única
Ucrânia	2011	2012
*URSS	1984	1987





















Quadro B: Países reconhecidos pela ONU (janeiro de 2016): grafia normalizada pelo VT

#	País	Bandeira
1	Afeganistão	
2	África do Sul	
3	Albânia	
4	Alemanha	
5	Andorra	
6	Angola	
7	Antígua e Barbuda	
8	Arábia Saudita	
9	Argélia	
10	Argentina	
11	Arménia / Armênia (Br.)	
12	Austrália	
13	Áustria	
14	Azerbaijão	
15	Baamas	
16	Bangladexe	
17	Barbados	
18	Barém	
19	Bélgica	
20	Belize	
21	Benim	
22	Bielorrússia / Belarus (Br.)	
23	Bolívia	
24	Bósnia e Herzegovina	
25	Botsuana	
26	Brasil	
27	Brunei	
28	Bulgária	
29	Burquina Fasso	
30	Burundi	
31	Butão	
32	Cabo Verde	
33	Camarões	
34	Camboja	
35	Canadá	
36	Catar	









37	Cazaquistão	
38	Chade	
39	Chile	
40	China	
41	Chipre	
42	Colômbia	
43	Comores	
44	Congo	
45	Coreia do Norte	
46	Coreia do Sul	
47	Costa do Marfim	
48	Costa Rica	
49	Croácia	
50	Cuaite (Hip.)	
51	Cuba	
52	Dinamarca	
53	Domínica / Dominica (Br.)	
54	Egito	
55	Emirados Árabes Unidos	
56	Equador	
57	Eritreia	
58	Eslováquia	
59	Eslovênia / Eslovênia (Br.)	
60	Espanha	
61	Estados Unidos	
62	Estônia / Estônia (Br.)	
63	Etiópia	
64	Fiji	
65	Filipinas	
66	Finlândia	
67	França	
68	Gabão	
69	Gâmbia	
70	Gana	
71	Geórgia	
72	Granada	
73	Grécia	
74	Guatemala	

75	Guiana	
76	Guiné	
77	Guiné-Bissau	
78	Guiné Equatorial	
79	Haiti	
80	Honduras	
81	Hungria	
82	Iémen / Iêmen (Br.)	
83	Ilhas Marshall	
84	Ilhas Salomão	
85	Índia	
86	Indonésia	
87	Irão / Irã (Br.)	
88	Iraque	
99	Irlanda	
90	Islândia	
91	Israel	
92	Itália	
93	Jamaica	
94	Japão	
95	Jibuti	
96	Jordânia	
97	Laos	
98	Lesoto	
99	Letónia / Letônia (Br.)	
100	Líbano	
101	Libéria	
102	Líbia	
103	Listenstaine	
104	Lituânia	
105	Luxemburgo	
106	Macedónia / Macedônia (Br.)	
107	Madagáscar / Madagascar (Br.)	
108	Malásia	
109	Maláui	
110	Maldívas	
111	Mali	
112	Malta	

113	Marrocos	
114	Maurícias / Maurício (Br.)	
115	Mauritânia	
116	México	
117	Myanmar / Birmânia	
118	Micronésia	
119	Moçambique	
120	Moldávia	
121	Mónaco / Mônaco (Br.)	
122	Mongólia	
123	Montenegro	
124	Namíbia	
125	Nauru	
126	Nepal	
127	Nicarágua	
128	Níger	
129	Nigéria	
130	Noruega	
131	Nova Zelândia	
132	Omã	
133	Países Baixos	
134	Palau	
135	Panamá	
136	Papua Nova Guiné	
137	Paquistão	
138	Paraguai	
139	Peru	
140	Polónia / Polónia (Br.)	
141	Portugal	
142	Quénia / Quênia (Br.)	
143	Quirguistão	
144	Quiribáti	
145	Reino Unido	
146	República Centro Africana	
147	República Checa / República Tcheca (Br.)	
148	República Democrática do Congo	
149	República Dominicana	
150	Roménia / Romênia (Br.)	

151	Ruanda	
152	Rússia	
153	Salvador	
154	Samoa	
155	Santa Lúcia	
156	Santa Sé	
157	São Cristóvão e Neves	
158	São Marino	
159	São Tomé e Príncipe	
160	São Vicente e Granadinas	
161	Seicheles	
162	Senegal	
163	Serra Leoa	
164	Sérvia	
165	Singapura / Cingapura (Br.)	
166	Síria	
167	Somália	
168	Seri Lanca (Hip.)	
169	Suazilândia	
170	Sudão	
171	Sudão do Sul	
172	Suécia	
173	Suíça	
174	Suriname	
175	Tailândia	
176	Tajiquistão	
177	Tanzânia	
178	Timor-Leste	
179	Togo	
180	Tonga	
181	Trindade e Tobago	
182	Tunísia	
183	Turquemenistão / Turcomenistão (Br.)	
184	Turquia	
185	Tuvalu	
186	Ucrânia	
187	Uganda	
188	Uruguai	

189	Usbequistão / Uzbequistão (Br.)	
190	Vanuatu	
191	Venezuela	
192	Vietname / Vietnã (Br.)	
193	Zâmbia	
194	Zimbábue	

Nota. «Br.» significa «forma exclusiva da variedade do Brasil».

Nota2. «Hip.» significa «forma ainda por validar».

Figura A: Placa toponímica de Armação de Pera, agosto de 2009



©2016 Google

Figura B: Placa toponímica para Armação de Pera, agosto de 2009



©2016 Google

Figura C: Placa toponímica de Troia, agosto de 2009



©2016 Google

Figura D: Placa toponímica de Troia, dezembro de 2014



©2016 Google



Figura E: Placa toponímica para Troia, agosto de 2009



©2016 Google

Figura F: Placa toponímica para Troia, janeiro de 2015



©2016 Google



Figura G: Placas toponímicas de Pero Pinheiro, setembro de 2009



©2016 Google

Figura H: Placas toponímicas de Pero Pinheiro, setembro de 2014



©2016 Google

Figura I: Placa toponímica de Crastos, novembro 2009



©2016 Google

Figura J: Placa toponímica de Crastos, fevereiro de 2015



©2016 Google