MINERAÇÃO DE TEXTO em Humanidades

Hélia Guerra | Luís Mendes Gomes | Mário Viana | Rosa Simas

Ficha Técnica

Edição

Centro de História da Universidade de Lisboa

Autores

Hélia Guerra
Luís Mendes Gomes
Mário Viana
Rosa Simas

Design & Paginação

DDDDproduções

Impressão

Nova Gráfica

Depósito Legal

397055/15

ISBN

978-989-8068-18-7

Índice

Prefácio 9

Introdução 15

Parte I – Elementos de programação em Python 19

1. Algoritmos e programação 21

Como aprender a programar 21

Instalar o Python 23

Olá, Python! 28

Efetuar cálculos 32

Guardar os programas em ficheiro 34

Interagir através do ecrã e do teclado 36

Desenvolvimento de programas 43

Fluxogramas 44

2. Conceitos e técnicas de programação 49

Tipos de dados elementares 49

O tipo inteiro (int) 51

O tipo real (float) 55

O tipo lógico (bool) 60

Variáveis e atribuição de valores 65

Documentação de programas 71

Estruturas de controlo 74

Seleção 76
Repetição 96
Cadeias de carateres (*strings*) 118
Listas 139
Dicionários 156
Ficheiros 168
Definição de funções 182
3. Expressões regulares 195
A construção de expressões regulares 196
Utilização do módulo Python RE 205
Bibliografia 227

Parte II – Aplicações a estudos de caso 229
4. Estudo de caso em História - Os casais despovoados nas inquirições régias portuguesas (1220-1258): uma aproximação com Python 231
O percurso 231
O problema 233
Uma aproximação com Python 236
script1H (criar dicionários) 237
script2H (corrigir *corpus*) 248
script3H (pesquisar por palavras) 252
script4H (pesquisar por prefixos) 257
script5_1H (contextos de palavra) 261

script5_2H (contextos de coocorrência) 266
script5_3H (contextos de prefixos) 269
script6_1H (densidade de ocorrências de palavra) 279
script6_2H (densidade de ocorrências de prefixo) 287
Bibliografia 293
5. Estudo de caso em Literatura - O romance *Mau tempo no Canal* de Vitorino Nemésio: abordagem com Python 295
O contexto 295
O desafio 297
Uma abordagem com Pyhon 230
A inventariação lexical 230
A seleção lexical 306
A contextualização 317
Bibliografia 330

Índice de Diagramas 332
Índice de Figuras 335
Índice de Programas 345
Índice de Tabelas 351

5. Estudo de caso em Literatura
O romance *Mau Tempo no Canal* de Vitorino Nemésio: uma abordagem com Python

O contexto

Em 1972, quando um dos autores deste livro foi estudar para a University of California, Los Angeles (UCLA), mal sabia que a Internet estava a nascer nos laboratórios desta universidade, ligada em rede com a University of California, Santa Barbara (UCSB), a Stanford University e a University of Utah. Filha de emigrantes luso-americanos, criada na Califórnia e desafiado, desde sempre, por duas línguas e culturas, foi para Los Angeles estudar português – língua, literatura e cultura. Com a intenção de complementar, assim, o inglês da sua formação escolar, e o espanhol que também dominava, posicionou-se naturalmente no estudo das Humanidades, que estavam alojadas no lado norte do campus da UCLA, enquanto, na parte sul, docentes e investigadores das Tecnologias desenhavam e montavam os alicerces da rede global que viria a ser a Internet.

Efetivamente, em 1961, no auge da Guerra Fria, a UCLA tinha recebido, do Department of Defense dos Estados Unidos, o Q-32, um daqueles computadores primitivos que eram batizados com um nome, em código, e que ocupavam uma sala inteira. Tratando-se de investigação, na altura ainda ligada a assuntos de defesa nacional, o sigilo seria grande. Assim, nos anos que esteve na UCLA, o Q-32 fazia parte de outro mundo, que não imaginava e ao qual não teve acesso. Isto, apesar de até ter frequentado o Departamento de Engenharia durante os dois anos em que, entre os horários das aulas, trabalhou em regime de *part-time* num estudo sobre mobilidade e meios de transporte. Porém tratava-se de mobilidade física – e não virtual – sobre terreno e em meios urbanos.

Mais de quatro décadas depois, a conjugação das Humanidades com as tecnologias aparece de novo no seu percurso, mas agora de uma forma concreta e ativa através da participação neste livro. Aceitou este desafio, como investigador das Humanidades sempre recetiva a novas abordagens e descobertas. Contudo, a sensação de ser "a stranger in a strange land" não lhe tem largado, levando-a a procurar informação sobre este novo "terreno", num esforço para conhecer as origens da Internet e perceber o desenvolvimento das novas tecnologias da informação e da comunicação. Nesta busca fora do que se poderá chamar a sua zona de conforto, aprendeu, por exemplo, que Vinton Cerf, agora conhecido como o "pai" da Internet, era também estudante universitário em 1972, ano do envio da primeira mensagem eletrónica. Numa perspetiva linguística, também percebeu que o próprio nome deste fenómeno tecnológico emergente teve várias versões, até que o termo "International Network" se convencionou na palavra "Internet" hoje ubíqua.

Veio a saber também que, passadas duas décadas, estando já a viver em Portugal, o protagonismo da era virtual tinha passado para a Europa. Efetivamente, em 1992, ano em que nasceu a sua filha, a *World Wide Web* (www) tinha igualmente o seu início, mas no seio do CERN (*Centre Européen de Recherche Nucléaire*), fruto da atividade do britânico Tim Berners-Lee, o putativo "pai" da www, e de outros investigadores deste Centro Europeu.

Aproximando da atualidade esta incursão pela cronologia dos avanços tecnológicos mais marcantes, percecionou a crescente velocidade com que as novidades vão aparecendo, tendência cada vez mais notável e marcante. O formato original da *World Wide Web*, de 1992, que globalizou o envio de emails e o uso dos motores de busca, deu lugar, em 2002, à Web 2.0, que possibilitou as partilhas interativas das redes sociais, dos blogues, dos wikis e dos chats. Assim, o *Facebook* aparece em 2004, o *YouTube* em 2005 e o *Twitter* em 2006. Logo de seguida, a Internet e a www entram na era móvel, com o *iPhone* em 2007 e depois com o *iPad*, em 2010. Estando a rede global da Internet e as potencialidades da www disponíveis nos *gadgets* móveis, nesse mesmo ano de 2010, surge a *Web* 3.0, um avanço do simples motor de busca para o que já foi apelidado de "motor de decisão", com potencialidades que se vão aproximando cada vez mais do que se convencionou chamar de inteligência artificial.

O desafio

Tendo integrado estes e outros dados no conhecimento exíguo que tem do mundo das Tecnologias, partiu para a concretização da sua participação neste projeto; isto é, avançou para a aprendizagem do que vem a ser outra linguagem – a do Python. Desde o princípio, achou que fazia sentido a aplicação de uma linguagem de programação para a mineração de texto, nas Humanidades, em geral, e na análise literária, especificamente. Nesse sentido, foi designado o texto a tratar, sendo consensual, entre os colegas da equipa, a escolha do romance de Vitorino Nemésio, *Mau Tempo no Canal*, considerado pela crítica como sendo a grande obra da literatura dos Açores e da Literatura Portuguesa do século XX. Como investigadora interessada no ramo da análise literária conhecido pelo nome de Ecocrítica, tencionava desenvolver um estudo focado na análise das representações da natureza dos Açores na grande obra de Vitorino Nemésio. No entanto, tinha a perfeita noção de que este seria um projeto de peso dadas as características do romance. Para além da própria dimensão do livro, com as suas quase quinhentas páginas, a obra recria dez localidades em quatro ilhas e reúne mais de cem personagens.

Aliada à complexidade do texto está o facto de esta ser a primeira vez que se estuda o romance com o suporte de ferramentas tecnológicas de mineração de texto. Efetivamente, para a análise de textos literárias, particularmente quando são multifacetados e extensos, a análise quantitativa e sistemática, pautando as potencialidades tecnológicas da linguagem de programação Python, surge como uma ajuda de grande utilidade. Concomitantemente, *Mau Tempo no Canal* nunca foi analisado à luz da ecocrítica, apesar de existirem alguns trabalhos inseridos em outros contextos que apontam nesse sentido.

Para além das abordagens científicas focando as questões ambientais no no contexto das ilhas dos Açores, como as *Crónicas do ambiente* de José Virgílio Cruz (2009), por exemplo, têm surgido alguns títulos no campo das humanidades. Na perspectiva dos estudos filosóficos sobre figuras do século

XIX, Magda Carvalho publicou *A natureza em Antero de Quental: o projecto de uma metafísica positivista* (2006) e José Brandão da Luz deliniou os "Percursos da mundividência naturalista nos Açores" (2009), no contexto novecentista, quando a filosofia positivista, "consagra a convicção inabalável de que o desenvolvimento das ciências emancipou a humanidade das dependências que a religião e a metafísica haviam estabelecido [e] contribuiu decisivamente para acreditar a mundividência naturalista que modelava muitos dos principais espíritos da ciência e da cultura da época" (p.14).

Já sob o olhar da ecocrítica, corrente de análise literária que vem à luz no fim da década de 1980, Vitor Mendes, no artigo "A ecocrítica e os Açores" (2009), propõe "a teorização da natureza" (156) à luz da cultura açoriana e do conceito de açorianidade, termo criado por Vitorino Nemésio em 1932. Neste contexo, desenvolvido um olhar ecocrítico sobre as representações da natureza açórica tecidas por americanos novecentistas em contato direto com as ilhas (Simas, 2014; 2013). Contudo, o olhar ecocrítico sobre *Mau Tempo no Canal* agora proposto configura-se como uma nova abordagem de um texto literário que foi amplamente estudado e comentado no passado.

Efetivamente, bastante antes de a corrente de análise literária ecocrítica ter surgido nos anos 80 do século XX, o grande romance de Vitorino Nemésio, três décadas após a sua publicação, em 1944, começa a merecer a atenção de uma variedade de críticos literários portugueses, incluindo os nomes conceituados da academia nacional. Numa breve síntese desses escritos, notamos que David Mourão-Ferreira (1974) abre caminho, ao classificar a obra como sendo simultaneamente um romance "realista e simbólico" (p. 126) que merece assinalável destaque no panorama da literatura portuguesa. Partindo da análise estrutural da obra, Maria Lúcia Lepecki (1974) diz que a construção de *Mau Tempo no Canal* assenta em dois eixos narrativos: um "horizontal" englobando o ambiente físico e a realidade social açoriana, e outro "vertical" abrangendo as várias fases do conflito amoroso vivido pela protagonista Margarida Clark Dulmo (p. 169). Ao assinalar os cinquenta anos da publicação original, a sétima edição do romance, de 1994, abre com uma Introdução de José Martins Garcia, que destaca o hábil manuseio da voz narrativa "filtrada"

pela personagem principal de Margarida, da seguinte forma:

> *A cumplicidade da voz narrativa com a personalidade de Margarida constitui, neste romance, um processo em que Nemésio depôs toda a sua capacidade de captação do real exterior e interior, bem como a sua exímia arte da palavra. Com efeito, Margarida é a personagem que mais profundamente apreende a envolvência insular, os efeitos da atmosfera sobre os nervos, os efeitos da rotina sobre a ansiedade, o impulso e os limites da evasão, o devaneio e o sonho, mas também a condição da mulher, os preconceitos de classe, a decadência presunçosa da casta, o dilema que se instala entre a busca de felicidade e a frustração. (15-16)*

Também focando a protagonista central, Heraldo da Silva (1985) descreve a multifacetada realidade açoriana do romance: o solo de vulcões e "mistérios" de lava que Margarida visita no Pico e em São Jorge; o mar real e mítico de barcos e baleias; o canal de tempestades e bonanças; a fauna e flora com cenas pastoris de ordenha de vacas e cavalgadas desportivas; as tradições e festividades populares com bordados e Festas do Espírito Santo; as origens históricas do povoamento das ilhas açorianas; e o desenvolvimento social e económico evocado nos antepassados de Margarida Clark Dulmo, sendo o Clark de origem saxónica e burguesa, e o Dulmo flamengo e nobre (288-289). Muito simplesmente, segundo João Gaspar Simões (1974), Margarida retrata as raízes genealógicas e míticas da sua terra natal, ao mesmo tempo que "aparenta a própria terra açoriana que a viu nascer" (p. 77), observações que apontam para o papel fulcral da protagonista, neta de Margarida Terra, nos múltiplos retratos da realidade açoriana que o romance apresenta. Neste contexto, começa-se a perceber que Margarida Clark Dulmo, a "pérola do Faial" (Nemésio, 1992: 61) – a "pérola nemesiana" (Simas, 1994) – está intimamente ligada à forma como o autor retrata a natureza açoriana no romance *Mau Tempo no Canal*.

Uma abordagem com Python

Nesta secção utilizamos a linguagem de programação Python para efetuar uma análise quantitativa ao texto literário *Mau Tempo no Canal*, que já vai em mais de dez edições, desde a sua publicação em 1944. Alguns dos scripts desenvolvidos no capítulo anterior são aqui reutilizados ou adaptados ao que se pretende estudar, nomeadamente para criar dicionários, pesquisar por palavras, obter contextos de palavras. Dadas as caraterísticas do texto, foi acrescentado um novo *script* para pesquisar notas de rodapé.

À semelhança do que fizemos no capítulo anterior, cada *script* é apresentado com um breve resumo, um diagrama explicativo das principais funções utilizadas, blocos de código e exemplos da sua execução e da janela interativa do Python (*shell*). A condição *sine qua non* é dispor do romance em ficheiros de texto (identificados pela extensão .txt).

A inventariação lexical

Para proceder à mineração e subsequente análise deste texto literário na íntegra, começou-se por inventariar as palavras escolhidas pelo autor para dar corpo à sua narrativa emblemática da experiência açoriana. Assim, logo à partida, a mineração de texto produz resultados interessantes referentes aos grupos lexicais e às funções das palavras utilizadas no livro.

Para obter um dicionário de vocábulos por ordem alfabética, criou-se o *script* `script1L`, tendo-se seguido os mesmos passos do `script1H` (*vide* Programa 4.1) para a geração deste tipo de dicionário. O *script* carrega os ficheiros de texto que se encontram no diretório atribuído pelo utilizador e, a partir deles, vai gerar o dicionário pretendido das palavras do corpus e guardar o resultado em ficheiro de texto e em ficheiro csv (*comma separated values*). O Diagrama 5.1 e o Programa 5.1 mostram, respetivamente, a arquitetura e o código do *script*. Recordamos que o procedimento consiste em converter

os ficheiros de texto correspondentes ao *corpus* numa única *string* que posteriormente é alterada por remoção de todos os carateres especiais que não façam parte do alfabeto das palavras. A lista de palavras do texto é obtida a partir desta *string* alterada. O utilizador pode intervir, indicando se pretende considerar a capitalização, bem como o comprimento mínimo das palavras a incluir no dicionário.

Diagrama 5.1 - `script1L` (criar dicionário por ordem alfabética).

```
def script1L(dir_corpus,dir_resultados):
    corpus=carregar_corpus(dir_corpus)
    print("Corpus carregado.")
    corpus_alterado=remover_caracteres(corpus)
    print("Corpus alterado.")
    lista=listar_palavras(corpus_alterado,minusculas=True)
    print("Lista de palavras gerada.")
    dic1=gerar_dicionario_chave_palavras(lista)
    print("Dicionário de palavras-frequencias gerado.")
    guardar_dicionario_palavras_ficheiro_texto(dic1,"todas_palavras",dir_resultados)
    print("Ficheiro de texto com dicionário de palavras-frequencias guardado com sucesso.")
    guardar_dicionario_palavras_ficheiro_csv(dic1,"todas_palavras",dir_resultados)
    print("Ficheiro csv com dicionário de palavras-frequencias guardado com sucesso.")

#-----------------------------------------------------------------------------------------

dir_corpus="C:\\Users\\MTH\\corpus\\mtc\\"
dir_resultados="C:\\Users\\MTH\\resultados\\mtc\\"
script1L(dir_corpus,dir_resultados)
```

Programa 5.1 - `script1L` (criar dicionário por ordem alfabética).

O *script* recorre a várias funções já definidas no capítulo anterior. O carregamento do *corpus* é efetuado através da função `carregar_corpus` (*vide* Programa 4.2). A remoção de carateres especiais é feita através da função `remover_carateres` (*vide* Programa 4.4) considerando a mesma lista de sinais. A função `listar_palavras` (*vide* Programa 4.7) cria a lista pretendida, ignorando a capitalização das palavras. A geração do dicionário considera todas as palavras e é feita através da função `gerar_dicionario_chave_palavra` (*vide* Programa 4.7). Os ficheiros de texto e no formato csv são criados, respetivamente, através das funções `guardar_dicionario_palavras_ficheiro_texto` (*vide* Programa 4.8) e `guardar_dicionario_palavras_ficheiro_csv` (*vide* Programa 4.10).

Na Figura 5.1 apresenta-se a shell que resultou da execução do *script* para o *corpus* do *Mau Tempo no Canal*, tendo sido criados dois ficheiros com o respetivo dicionário.

```
Corpus carregado.
Corpus alterado.
Lista de palavras gerada.
Dicionário de palavras-frequencias gerado.
Ficheiro de texto com dicionário de palavras-frequencias guardado com sucesso.
Ficheiro csv com dicionário de palavras-frequencias guardado com sucesso.
```

Figura 5.1 - *Shell* da execução do `script1L`.

Entre os dados obtidos, podemos sinalizar os que são mais significativos para a análise que se pretende. Por exemplo, para se proceder a uma abordagem ecocrítica de *Mau Tempo no Canal*, na Tabela 5.1 foram isoladas, mantendo a ordem alfabética, as palavras relacionadas com "natureza" e "natural", incluindo "nature" – um dos variados exemplos de vocábulos em inglês que Nemésio utiliza no romance. Neste grupo lexical, a forma singular e plural do adjetivo "natural" ocorre 20 vezes, enquanto o advérbio "naturalmente" aparece 13 vezes. Tendo esta informação, podemos construir um *script* para localizar estas palavras no texto e comparar os contextos e situações em que são utilizadas. O mesmo se poderá fazer com a grande variedade de palavras relacionadas com a atividade de navegação marítimo, através dos grupos lexicais,

também na Tabela 5.1, que se prendem com palavras como "navegar", "navegador", "navio" (a palavra mais usada deste grupo), etc., incluindo "navigateur" do francês, outra língua estrangeira presente no romance.

naturais	4	naus	1	navegar	3
natural	16	nauseada	1	navegasse	1
naturalidade	6	naval	3	navegava	1
naturalista	1	navalha	3	navegação	5
naturalmente	13	navalhas	1	naves	1
nature	1	nave	6	navigateur	1
natureza	4	navegador	1	navio	28
nau	3	navegadores	1	navio-escola	2
naufrágio	2	navegando	2	navios	6

Tabela 5.1 - Exemplo da variedade lexical de dois grupos de palavras.

Desta forma, começamos a notar as características e as tendências da variedade lexical do texto em apreço. Querendo reconfigurar e reavaliar a pertinência dos resultados já obtidos, podemos avançar para uma segunda inventariação das palavras do texto, mas agora pela ordem numérica da sua ocorrência, que pode ser perscrutada por ordem decrescente ou crescente. Para tal, recorremos a uma parte do `scriptlH` que permite gerar um dicionário ordenado pelo número de ocorrências, associando a cada valor a respetiva lista de palavras que verifica esse número de ocorrências. O diagrama do *script* é idêntico ao do `scriptlL`, cujo código se apresenta no Programa 5.2.

```
def script2L(dir_corpus,dir_resultados):
    corpus=carregar_corpus(dir_corpus)
    print("Corpus carregado.")
    corpus_alterado=remover_caracteres(corpus)
    print("Corpus alterado.")
    lista=listar_palavras(corpus_alterado,minusculas=True)
    print("Lista de palavras gerada.")
    dic1=gerar_dicionario_chave_palavras(lista)
    dic2=gerar_dicionario_chave_ocorrencias(lista)
    print("Dicionário de frequencias-palavras gerado.")
    guardar_dicionario_ocorrencias_ficheiro_texto(dic2,"ocorrencias_todas_palavras",dir_resultados)
    print("Ficheiro de texto com dicionário de frequencias-palavras guardado com sucesso.")
    guardar_dicionario_ocorrencias_ficheiro_csv(dic2,"ocorrencias_todas_palavras",dir_resultados)
    print("Ficheiro csv com dicionário de frequencias-palavras guardado com sucesso.")

#-----------------------------------------------------------------------------------------------

dir_corpus="C:\\Users\\MTH\\corpus\\mtc\\"
dir_resultados="C:\\Users\\MTH\\resultados\\mtc\\"
script2L(dir_corpus,dir_resultados)
```

Programa 5.2 - `script2L` para obter o dicionário de palavras por ordem numérica.

Neste contexto, notamos que a mineração de texto produz resultados consideráveis referentes às funções gramaticais dos vocábulos que compõem o romance (*vide* Tabela 5.2 e Tabela 5.3). Como seria de esperar, predominam certos tipos de palavras devido às suas funções linguísticas, intrínsecas ao funcionamento da língua. As hipóteses de análise são muitas e variáveis, dependendo dos dados que se pretende e dos objetivos que se almeja.

6102	de	1481	uma	960	no
5052	a	1295	não	901	as
4096	e	1246	com	884	ao
3972	o	1097	para	850	em
3332	que	1053	os	752	como
2573	do	1042	à	725	é
2333	da	1035	na	670	das
1984	um	982	se	662	**Margarida**

Tabela 5.2 - Palavras mais frequentes e número de ocorrências.

662	**Margarida**	400	mais	314	lhe
653	dos	365	seu	298	casa
628	mas	360	lá	293	ou
591	por	356	sua	291	já, **Roberto**
569	era	**336**	**João**	**285**	**tio**

Tabela 5.3 - Nomes próprios mais frequentes e número de ocorrências.

Entre as palavras mais frequentes (Tabela 5.2), por exemplo, os artigos e as preposições predominam. Nas suas formas variadas, a preposição "de" é dominante: "de" aparece 6 102 vezes, "do" 2 573, "da" 2 333 e "das" 670 (Tabela 5.2), enquanto "dos" figura 653 vezes, logo depois de "margarida" (Tabela 5.3). Também dominam os artigos definidos e indefinidos, igualmente nas suas diversas formas. Focando de novo as preposições, notamos que, por si só, a preposição "para" surge 1097 vezes, enquanto as várias configurações de "em" também sobressaem claramente: "na" aparece 1 035 vezes, "no" 960, "em" 850, "nos" 263 e "nas" 261 vezes. A conjunção "e" é a terceira palavra mais usada no romance, num total de 4 096 vezes, (Tabela 5.3), enquanto os adjetivos possessivos são também frequentes, especialmente nas suas formas singulares: "seu" surge 365 vezes e "sua" 356 vezes (Tabela 5.3).

Indicativo, logo à partida, do papel preponderante da personagem principal do romance, o nome de Margarida é a vigésima-quarta palavra mais frequente do romance, figurando entre a listagem da Tabela 5.2, que contém os tipos de vocábulos estruturantes que desempenham as funções essenciais ao funcionamento gramatical e sintáctico desta e de qualquer língua.

Seguindo ainda a sequência dos vocábulos mais frequentes, mas passando agora para os nomes próprios dos protagonistas do romance emblemático dos Açores, não surpreende que o nome da personagem principal surja num total de 699 vezes – 662 como Margarida (Tabela 5.3) e, entre os termos menos usados, 33 na forma popular e afetuosa de "Bidinha" e 4 como "Bida". No contexto dos nomes próprios mais frequentes (Tabela 5.2), também como seria de esperar, encontramos as referências às duas personagens masculinas

que contracenam com Margarida ao longo do romance – o nome do pretendente "João", filho da família Garcia, aparece 336 vezes, seguido de "Roberto", o tio de Margarida que se encontra de visita da Inglaterra no decorrer da ação do romance, personagem cujo nome de batismo surge em 291 ocasiões, enquanto o título "tio" é referido 285 vezes.

A seleção lexical

Os *scripts* e os resultados apresentados até aqui incidem sobre um texto no seu todo, sem seleção alguma: o `script1L` leva à ordem alfabética de todos os dados, enquanto o `script2L` origina resultados ordenados por ordem numérica. Como já tivemos oportunidade de ver no capítulo 4, muitas vezes, há situações em que se pretende a obtenção de resultados selecionados de acordo com critérios previamente estabelecidos. Para exemplificar, o foco mantém-se nos nomes pessoais do romance, mas passamos agora para os apelidos mais significativos. Neste contexto, realizamos uma pesquisa seletiva dos nomes de família mais frequentes do texto – Dulmo e Clark pela parte de Margarida, e Garcia do pretendente João. Para tal, recorremos ao `script3H` que aqui renomeamos para `script3L` apenas para estar alinhado com os restantes *scripts* do capítulo e facilitar a tarefa ao leitor que apenas tenha interesse neste capítulo. A sua arquitetura encontra-se no Diagrama 5.2 e o respetivo código no Programa 5.3.

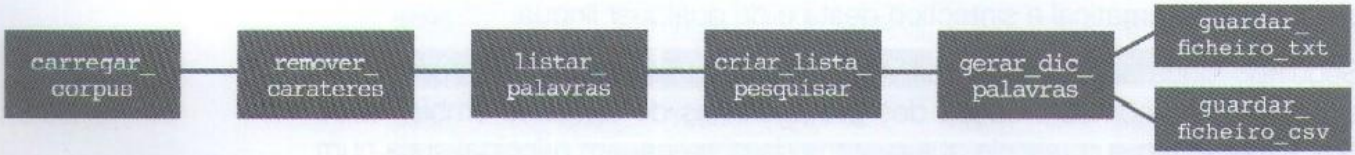


Diagrama 5.2 - `script3L` (pesquisar por palavras).

```
def script3L(dir_corpus,dir_resultados):
    corpus=carregar_corpus(dir_corpus)
    corpus_alterado=remover_caracteres(corpus)
    lista_pesquisar=criar_lista_palavras_pesquisar()
    print("Lista de palavras a pesquisar")
    print(lista_pesquisar)
    cap=input("Ignorar a capitalização? (1 - sim; 0 - não) ")
    if int(cap)==1:
        minusculas=True
    else:
        minusculas=False
    lista=listar_palavras(corpus_alterado,minusculas)
    dic_base=gerar_dicionario_chave_palavras(lista)
    dic=gerar_dicionario_frequencias_lista_pesquisar(dic_base,lista_pesquisar)
    f_txt=input("Guardar resultados em ficheiro de texto? (1 - sim, 0- não) ")
    if int(f_txt)==1:
        nome_ficheiro=input("Indique o nome do ficheiro: ")
        guardar_frequencias_ficheiro_texto(dic,nome_ficheiro,dir_resultados)
        print("Ficheiro guardado com sucesso.")
    f_csv=input("Guardar resultados em ficheiro csv? (1 - sim, 0- não) ")
    if int(f_csv)==1:
        nome_ficheiro=input("Indique o nome do ficheiro: ")
        guardar_dicionario_palavras_ficheiro_csv(dic,nome_ficheiro,dir_resultados)
        print("Ficheiro guardado com sucesso.")
#-----------------------------------------------------------------------------------------

dir_corpus="C:\\Users\\MTH\\corpus\\mtc\\"
dir_resultados="C:\\Users\\MTH\\resultados\\mtc\\"
script3L(dir_corpus,dir_resultados)
```

Programa 5.3 - `script3L` (pesquisar por palavras).

Este *script* recorre a várias funções já definidas no capítulo anterior. A função **`criar_lista_palavras`** irá criar a lista com as palavras indicadas pelo utilizador (*vide* Programa 5.7). Segue-se a função **`gerar_dicionario_palavras_chave`** que irá criar o dicionário das palavras pretendidas (*vide* Programa 5.6).

A execução do **`script3L`**, neste caso para a seleção dos nomes de família pretendidos, encontra-se na Figura 5.2. Este *script* pode ser adaptado para outras pesquisas, tais como a pesquisa de termos compostos, como Clark-Dulmo, o apelido composto que está igualmente presente no romance, mas com menor expressão.

```
>>>
Introduza uma palavra: garcia
Pretende introduzir mais palavras? (1 - Sim, 0 - Não) 1
Introduza uma palavra: dulmo
Pretende introduzir mais palavras? (1 - Sim, 0 - Não) 1
Introduza uma palavra: clark
Pretende introduzir mais palavras? (1 - Sim, 0 - Não) 1
Introduza uma palavra: garcias
Pretende introduzir mais palavras? (1 - Sim, 0 - Não) 1
Introduza uma palavra: dulmos
Pretende introduzir mais palavras? (1 - Sim, 0 - Não) 1
Introduza uma palavra: clarks
Pretende introduzir mais palavras? (1 - Sim, 0 - Não) 0
Lista de palavras a pesquisar
['clark', 'clarks', 'dulmo', 'dulmos', 'garcia', 'garcias' 1
Ignorar a capitalização? (1 - sim; 0 - não) 1
Guardar resultados em ficheiro de texto? (1 - sim, 0- não) 1
Indique o nome do ficheiro: apelidos
Ficheiro guardado com sucesso.
Guardar resultados em ficheiro csv? (1 - sim, 0- não) 1
Indique o nome do ficheiro: apelidos
Ficheiro guardado com sucesso.
>>>
```

Figura 5.2 – *Shell* de execução do `script3L` para selecionar nomes de família.

Por sua vez, os resultados obtidos nesta execução do *script* vêm ilustrados a seguir, na Tabela 5.4, onde as ocorrências dos nomes de família mais significativos estão organizados por ordem alfabética. Em todo o caso, a pesquisa incide nem só sobre as formas singulares dos sobrenomes mencionados, que são mais frequentes, mas também sobre as formas plurais dos mesmos que, apesar de serem menos expressivas, são indicativas do papel fundamental, no decorrer da narrativa, do sentimento de pertença, ou não, a um destes grupos familiares, clãs que se identificam como tal, muitas vezes agindo e reagindo em conjunto, uns com ou contra os outros.

Clark	77	Dulmo	198	Garcia	250
Clarks	27	Dulmos	29	Garcias	22

Tabela 5.4 - Ocorrências dos nomes de família mais significativos.

Entre os nomes de família do elenco que compõe o romance em estudo, seria de esperar que figurasse mais o sobrenome da família de Margarida, "Dulmo" (empregue 198 vezes). Porém esse lugar cabe ao sobrenome da família "Garcia" (utilizado 250 vezes), como constatamos na Tabela 6.4. Esta preponderância aponta para a importância dada, ao longo de toda a narrativa, à grande rivalidade entre o pai de João, Januário Garcia, personificação da burguesia, então em franca acensão, e o pai de Margarida, Diogo Dulmo, patriarca representativo da classe aristocrática, em processo de decadência lenta mas inevitável. Protagonistas do conflito que molda o romance, ambos estão inseridos numa sociedade de cariz marcadamente tradicional, confrontada com os transtornos e as intempéries que caracterizam os períodos de transformação social.

Antes de passar para a forma de situar a pesquisa de ocorrências no contexto do discurso narrativo do texto, achamos importante exemplificar outros resultados da busca seletiva que merecem ser referidos no âmbito da análise de *Mau Tempo no Canal*. Para estudar a riqueza lexical que Nemésio imprime ao seu romance, por exemplo, torna-se evidente o grande desafio que é estudar a linguagem de todo o texto e analisar a diversidade dos modos de falar das múltiplas personagens. Neste contexto, sobressai a pujança do falar popular recriado pelo autor, que pode ser amplamente inventariado e estudado recorrendo à linguagem de programação Python.

Aliás, o uso deste código de programação é ideal para a análise linguística, especialmente de um texto complexo e multifacetado como o grande romance que Nemésio nos legou. Neste âmbito, a linguagem oral merece especial atenção, pela sua variedade e veracidade. Entre as características dos diálogos criados pelo narrador, por exemplo, o falar popular dos Açores sobressai ao longo do livro inteiro. Para estudar esta dimensão do romance, voltamos a implementar o `scriptlL` que, lembramos, é utilizado para a busca alfabética e integral de todos os vocábulos que compõem um texto. Assim, é possível chegar aos resultados isolados na Tabela 5.5, que exemplificam as diversas formas como a palavra "homem" é utilizada e pronunciada ao longo do romance, com destaque para o uso popular de "home" (plural de "homes").

home	16	homenagem	3	homenzinho	1
homem	67	homens	1	homes	2

Tabela 5.5 - Exemplos do falar popular (obtido com o `scriptlL`).

Efetivamente, o discurso oral nemesiano prende-se com variados registos e aspetos sociolinguísticos, e merece particular atenção, podendo ser estudados de múltiplas perspetivas e formas. Numa expressão da realidade açoriana, especialmente no Faial/Pico, no início do século XX, a voz narrativa conjuga dois mundos, o campestre e o cosmopolita. Por exemplo, a Figura 5.3, tirada da página 311 da edição de 2001 do romance, ilustra a forma como o autor combina vários registos linguísticos, neste caso, o falar popular e vocábulos derivados do Inglês. Aqui, no falar do antigo baleeiro e trancador Amaro de Mirateca, em vez da forma padrão – "homem" – do português, figura o uso de "home", observado na Tabela 5.5, entre outros usos característicos da pronúncia popular açoriana. Surge também o uso de termos derivados do inglês, incluindo a pronúncia popular – "charape!" – do comando informal "shut up!", explicado pelo autor na nota de rodapé na página, como se pode constatar na Figura 5.3.

dou o navio à América para meter aparelho. O ti Amaro, pelos vistos, pertenceu à tripulação... O *Cidade da Horta* começou a balear em Julho.

O velho trancador abanava a cabeça, pasmado:

– Como Vossa Sinhoria dá rezão disso tudo! Causos de há sessent' anos!... – Margarida abria docemente os olhos em direcção ao tio; Diogo Dulmo, sempre silencioso, deu volta à torneira de uma pipa e encheu dois copos. O velho retomou o fio interrompido: – Foi antão que m'assucedero os causos que stive aqui a cuntar ò titio e ò papai... cumprende a menina? Passei as passas do Algarve... penei mistério! Mágoas tamanhas...! – Margarida ainda pensou em pedir ao ti Amaro que repetisse «os causos»; mas sentia-se ausente, como que encantada, e o seu sorriso sonâmbulo entrava nos olhos do velho com a precisão e o efeito de uma abelha na flor. Ti Amaro, pegando enfim no copo e cheirando o vinho, resumiu: – Pena-se munto nesses mares, mãis aprende-se mais que nũa iscola! Eu andava por ì rotinho, e minha mãe passou trabalhos pra m'introixar a roipa. Roipa?... O quê!...: uns bidalhos! É verdade que suei mais d'um ano a atirá' lenha ao *try-work,* a pé do mastro da misena. Cada caldeiro daqueles lovava 36 galães d'azeite! As banhas da baleia mitio ânsias de lançar... Aquêl fartume a torresmo! Lá m'asservei...

«Um dia o Capitão chamou-me à ré e disse: "Pórigui! Falta-me um home ò *tub,* charape! Vais tu...".[1] Dipois passei a um patacho aqui no Faial, chomado *Feiticeira dos Mares* (o alma do diabo tinha mêmo feitiço!), e tranquei a minha promeira baleia, um cachalote de noventa barris, que nã se virava im seco!

«A *Gazela* aparlhou prà nossa (i) arte, e eu, que já stava casado e cheio de meninos e tinha minha mãe às costas (o patacho apanhava pouco... era a modo cobranto do nome!), lá fui atrás d'ũa sòldada maior. Aquilo sim! Bum

[1] *Portuguese. Tub*: um dos cinco lugares aos remos das canoas americanas: *harpooner, bow, midship, tub, apter-oar*, ou seja: trancador, 1.º remador, remo de meia-nau, *baquet*, remo de popa. «Charape»: ingl. *shut up!*

311

Figura 5.3 - Imagem da página 311 da edição de 2001 de *Mau Tempo no Canal.*

Como foi referido e se pode comprovar na página do livro reproduzida na Figura 5.3, os estrangeirismos fazem parte do romance, com a especial preponderância de neologismos derivados do inglês, expressão viva da ligação do arquipélago com os Estados Unidos, especialmente, mas também com a Inglaterra (na pessoa do Tio Roberto), tendência do falar dos Açores que Nemésio teve o extremo cuidado de recriar em contextos variados e de formas multifacetadas. Exemplos abundam e o tratamento varia. Por vezes, expressões em inglês aparecem autonomamente no discurso do romance, enquanto noutras ocasiões são explicadas no próprio texto, como acontece na página 117, com "alvarozes", pronúncia popular explicada por Margarida – "Uma espécie de fato de macaco, *over all*", ao que o tio Roberto responde "*Yes, I know...*" Noutras alturas, termos comuns no falar açoriano, adaptações do inglês à pronúncia portuguesa, também figuram sem explicação, como acontece com "calafona" (pp. 323 e 425), termo popular derivado de "California" que refere pessoas dos Açores que estão emigradas, ou no próprio estado da Califórnia ou simplesmente na América do Norte. O mesmo acontece quando um baleeiro interpela Margarida, dizendo "A menina nã vê?...o mar "tá rofe" (p. 400), numa apropriação, sem explicação, do adjetivo "rough" em inglês, apesar de este uso ser explicado em nota de roda pé na página 326, como consta à frente na entrada 12 da Tabela 5.6. De facto, a língua inglesa perpassa todo o romance de tal forma que, perto do fim, o narrador recorre a um termo em inglês – "Azorean torpor" – para fazer o diagnóstico do impacto do clima na população das ilhas – não obstante o adjetivo "Azorean" ser escrito com letra maiúscula em inglês, e não como aparece abaixo, na citação referida:

> *Um céu de algodão sujo tolda o arquipélago das nove ilhas: o 'mormaço' apaga os contornos do mar e da terra, e, amolecendo os pastos à custa da pele do proprietário e do pastor, dilui e arrasta as vontades, dá a homens e a coisas uma doença quase de alma, a que os ingleses, médicos do bem-estar, puseram uma etiqueta como quem descobre uma planta nova neste mundo seco e velho: azorean torpor. (p. 422)*

Noutras ocasiões, para superar o eventual desconhecimento, entre os seus leitores, dos termos vindos do Inglês, e também de outros usos e costumes característicos dos Açores, Vitorino Nemésio recorre a notas de rodapé, como já constatamos na Figura 5.3. Para pesquisar uma característica específica do texto, neste caso as notas de roda pé referindo termos derivados do Inglês, desenvolvemos o `script4L`.

A arquitetura do *script* encontra-se no Diagrama 5.3. Recorremos às expressões regulares para pesquisar ocorrências de frases precedidas com um número no início da linha, que consideramos ser o formato mais comum para a escrita das notas de rodapé. Naturalmente que como resultado poderá surgir alguma frase do texto cujo inicio da linha também seja um número. Neste caso, cabe ao utilizador, selecionar manualmente as notas de rodapé, uma vez que o texto não tem qualquer anotação que permita ao *script* reconhecer as ocorrências que efetivamente correspondem a notas de rodapé. Note-se que o processo de anotação de texto é um assunto não abordado no âmbito deste livro.

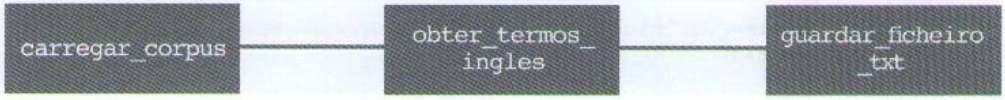


Diagrama 5.3 - `script4L` (pesquisar por notas de rodapé).

O código do *script* encontra-se no Programa 5.4. A função `obter_termos_ingles` (*vide* Programa 5.5) é que irá pesquisar as ocorrências no texto com o formato que definimos para encontrar as notas de rodapé. Este baseou-se no facto do autor utilizar a indicação "Ingl." ou "ingl." logo antes da anotação da grafia da língua inglesa.

```
def script4L(dir_corpus,dir_resultados):
    corpus=carregar_corpus(dir_corpus)
    contextos=obter_notas_termos_ingles(corpus)
    if contextos:
        nome_ficheiro=input("Nome que pretende atribuir ao ficheiro: ")
        guardar_texto_ficheiro(contextos,nome_ficheiro,dir_resultados)
        print("Ficheiro guardado com sucesso.")

#-----------------------------------------------------------------------------

dir_corpus="C:\\Users\\MTH\\corpus\\mtc\\"
dir_resultados="C:\\Users\\MTH\\resultados\\mtc\\"
script4L(dir_corpus,dir_resultados)
```

Programa 5.4 - `script4L` para selecionar notas de rodapé referindo o Inglês.

```
import re

def obter_termos_ingles(texto):
    er="[\s]"+"Ingl."+"(\s)(.)*?[\.\n]"
    m=re.finditer(er,texto,flags=re.IGNORECASE)
    freq=0
    contextos="********** Contextos *****************\n"
    for i in m:
        contextos+="\n"+i.group().strip()
        freq=freq+1
    if freq>0:
        return "\nTotal: "+str(freq)+" ocorrências.\n\n"+contextos
```

Programa 5.5 - Obter notas de rodapé em inglês.

Na Figura 5.4 apresentamos a *shell* resultante de uma execução do `script4L`.

```
>>>
Nome que pretende atribuir ao ficheiro: termos_ingles
Ficheiro guardado com sucesso.
>>>
```

Figura 5.4 - *Shell* da execução do `script4L`.

Os resultados da inventariação dos termos do inglês em notas de rodapé vêm elencados na Tabela 5.6. Arrematando esta discussão do uso do inglês no romance nemesiano, convém notar que, quer surjam integradas no discurso narrativo ou explicadas em nota de rodapé, por vezes estes usos e referências contém erros, especialmente na grafia, como tem vindo a ser comentado (*e. g.*, Fagundes, 1998). É este o caso quando o autor indica, em nota de roda pé, que a pronúncia popular "Arioche" (p. 287) refere o oceano Glacial Árctico, escrevendo "Artic ocean" (Tabela 5.6) em vez de "Arctic ocean", a grafia correta em inglês. De notar, na Tabela 5.6, a nota 6 referente ao termo inglês "*shut-up*", ilustrado na Figura 5.3. Efetivamente, para o estudo cabal de um texto com a variedade lexical e a densidade semântica do romance *Mau Tempo no Canal*, a programação Python surge como uma ferramenta ideal, contribuindo, de uma forma prática e sistemática, para explorar esta dimensão do texto, que ilustra as múltiplas ligações da cultura e a linguagem dos Açores, com a realidade anglo-americana.

1	1 Ingl. Artic ocean, oceano Glacial Árctico. (p. 287)
2	2 Ingl. Western Ground, Mar das Antilhas. (p. 287)
3	3 Ingl. Japanese, Japoneses (Mares do Japão). (p. 287)
4	1 Ingl: money. (p. 288)
5	1 Ingl: Nantucket. (p. 309)
6	1 Portuguese. Tub: um dos cinco lugares aos remos das canoas americanas: barpooner, bow, midsbip, tub, apter-oar, ou seja: trancador, 1.° remador, remo de meia-nau, banquet, remo de popa. «Charape»: ingl. shut up! (p. 311)
7	1 Ing.: Spade, espada. (p. 312)
8	1 Ingl: market – mercado (p. 315)
9	1 Ingl.: «She blows»; «ela esguicha». falando-se do jacto da baleia. (p. 320)
10	1 Ingl.: Frock: espécie de blusa de homem. (p. 321)
11	2 Ingl.: Pull ahead!: «puxa!». (p. 323)
12	1 Ingl.: Rough, «áspero», «agitado». (p. 326)

13	1 Ingl.: Log head: cepo cilíndrico fixado verticalmente no leito da proa das canoas baleeiras, em torno do qual corre a linha presa ao arpão enterrado no cachalote. (p. 327)
14	1 Ingl.: Spades: instrumentos cortantes empregados para decepar as baleias. (p. 332)
15	2 Ingl.: Saioing-machine (P): cutelo com dois cabos no prolongamento das costas da lâmina. (p.332)
16	2 Ingl.: Handle biller: balde de ferro, de cabo comprido, empregado na recolha do óleo de cachalote. (p. 335)
17	1 Ingl.: Indian rubber: borracha ou cauchu; por extensão, botas de cauchu. (p. 340)
18	1 Ingl.: Sweater. (p. 350)

Tabela 5.6 - Exemplos dos termos do inglês em notas de rodapé.

Convém registar que este *script* poderá ser facilmente adaptado para outras pesquisas, como por exemplo, notas de rodapé em geral. No Programa 5.6 apresentamos uma função que poderá ser usada.

```
import re

def obter_notas_rodape(texto):
    er="[\n]"+"[\d](\s)(.)*[\.\n]"
    m=re.finditer(er,texto,flags=re.IGNORECASE)
    freq=0
    contextos="********** Contextos *****************\n"
    for i in m:
        contextos+="\n\n"+i.group().strip()
        freq=freq+1
    if freq>0:
        return "\nTotal: "+str(freq)+" ocorrências.\n\n"+contextos
```

Programa 5.6 - Obter notas de rodapé.

```
def script4L(dir_corpus,dir_resultados):
    corpus=carregar_corpus(dir_corpus)
    contextos=obter_notas_rodape(corpus)
    if contextos:
        nome_ficheiro=input("Nome que pretende atribuir ao ficheiro: ")
        guardar_texto_ficheiro(contextos,nome_ficheiro,dir_resultados)
        print("Ficheiro guardado com sucesso.")
```

Programa 5.7 - `script4L` para selecionar notas de rodapé.

Por fim, vamos partir da referência à palavra "oceano" para explorar mais uma valência da utilização da programação Python que é fundamental para a análise literária: a pesquisa do(s) contexto(s) em que surge uma palavra ou sintagma. Antes, porém, voltamos à palavra "home" para ilustrar a importância da contextualização na análise discursiva e literária.

A contextualização

No seguimento dos *scripts* já ilustrados, falta indicar como se prossegue à pesquisa do contexto em que um dado termo é utilizado no texto. Para situar o sintagma em apreço no decorrer da narrativa, é possível configurar a programação de forma a moldar os resultados de acordo com o que antecede e segue um dado termo. Esta contextualização é essencial para se poder, com base no uso que é feito das palavras que compõe um dado texto, avaliar e analisar o mesmo, com especial incidência na sua significância literária.

Para ilustrar, voltemos ao termo "home", já referido na Tabela 5.5. Efetivamente, como já foi referido, o uso deste vocábulo poderá indicar a pronúncia popular da palavra "homem" em português. No entanto, num romance como *Mau Tempo no Canal*, onde o inglês está bem presente, "home" poderá igualmente referir a palavra "casa" na língua inglesa. Como é evidente, o sentido que o termo assume depende do encadeamento e da forma como é utilizado.

Recorremos ao `script5_1H` do capitulo anterior para efetuar as pesquisas pretendidas. À semelhança do que fizemos para o `script3L`, vamos renomear este *script* no contexto deste capítulo para `script5L`. A respetiva arquitetura encontra-se descrita no Diagrama 5.4. É efetuada uma pesquisa às ocorrências de uma palavra indicada pelo utilizador, obtendo-se um extrato do texto onde se encontra cada ocorrência que é guardado num ficheiro de texto. Mais uma vez, o uso de expressões regulares foi importante para caraterizar as pesquisas a realizar sobre o texto e obter os respetivos contextos.

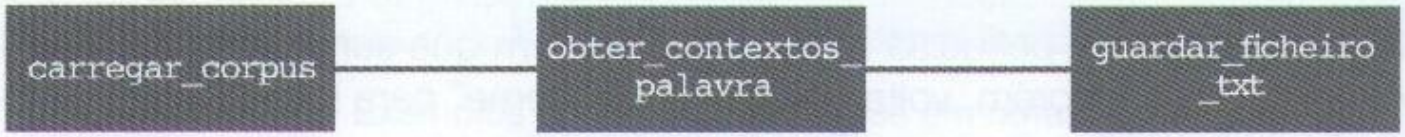


Diagrama 5.4 - `script5L` (contextos de palavra).

No Programa 5.8 encontra-se o código deste *script*. As funções que o integram já foram descritas no capítulo anterior (*vide* Programa 4.55 e Programa 4.56). O utilizador deverá indicar um valor para a dimensão do contexto pretendido em número de carateres. Esse valor servirá para calcular a posição mais à esquerda do contexto, através da sua subtração ao índice do primeiro carater da ocorrência da expressão regular. Se o resultado for negativo, considera 0 que corresponde ao índice do primeiro carater do texto; se não, considera o valor do resultado ou o menor valor mais próximo do resultado, caso o carater correspondente ao resultado seja alfanumérico. O cálculo da posição mais à direita é feito de forma análoga, adicionando o valor indicado ao índice do carater mais à direita da ocorrência da expressão regular no texto. Considerar o índice do último carater do texto, se o resultado for superior a este, caso contrário considerar o resultado ou o menor valor superior a este que corresponda a um carater não alfanumérico.

```
def script5L(dir_corpus,dir_resultados):
    corpus=carregar_corpus(dir_corpus)
    palavra=input("Palavra a pesquisar: ")
    cap=input("Ignorar a capitalização? (S - 1 N - 0) ")
    limInf=input("Número de caracteres à esquerda: ")
    limSup=input("Número de caracteres à direita: ")
    if int(cap)==1:
        contextos=obter_contextos_palavra(palavra,corpus,int(limInf),int(limSup),True)
    else:
        contextos=obter_contextos_palavra(palavra,corpus,int(limInf),int(limSup))
    if contextos:
        nome_ficheiro=input("Nome que pretende atribuir ao ficheiro: ")
        guardar_texto_ficheiro(contextos,nome_ficheiro,dir_resultados)
        print("Ficheiro guardado com sucesso.")

#-----------------------------------------------------------------------------------

dir_corpus="C:\\Users\\MTH\\corpus\\mtc\\"
dir_resultados="C:\\Users\\MTH\\resultados\\mtc\\"
script5L(dir_corpus,dir_resultados)
```

Programa 5.8 - `script5L` (pesquisar contextos de palavra).

O `script5L` foi executado para se conhecer o contexto da palavra "home" (*vide* Figura 5.5).

```
>>>
Palavra a pesquisar: home
Ignorar a capitalização? (S - 1 N - 0) 1
Número de caracteres à esquerda: 50
Número de caracteres à direita: 50
Nome que pretende atribuir ao ficheiro: home50
Ficheiro guardado com sucesso.
>>>
```

Figura 5.5 - *Shell* da execução do `script5L` para a palavra "home".

Na Tabela 5.6 apresentamos o contexto das 16 ocorrências encontradas para esta palavra. Comprovamos que, apesar de "home" como pronúncia popular de "homem" predominar claramente, o vocábulo em inglês também ocorre, referindo as raízes inglesas da família Clark (ocorrência 1, na Tabela 5.6).

1	fechavam através das gerações sobre os Clarks fiéis ao home, os Clarks de Ladlooke Grove, antes da hora da morte. (p. 55)
2	sangrou, se a foi ferrar? Assim é que me zelas as coisas? .. - Home, essa, patrão! - É só isto que eu digo: porque é que (p.94)
3	co a velhinha! Rapariga nã te fies im palavras de home rico: São como calhau rolado. Qu' anda na costa do (p. 131)
4	irmã é ila prove de Cristo; inviuvou há um ano. O home deixou-le um alqueirinho de terra e a casa: um palheirinho (p. 158)
5	inda que queiro, não atímum ... O campo carece mão de home. - Bem; vamos a ver ... - João Garcia puxou de lápis e (p. 159)
6	- ora a alembrança da menina! Fazer pouco de um home ... Um velho, cos dentes escabaçados! - E alargava a mão (p. 212)
7	é isso?! - Toca no velhote, Intavante! Toca, s'és home! - Caluda! Haja tenência ... o João da Cezilha, de chapéu (p. 284)
8	... Era um intretém ... (Risota.) - Dêixim lá o home! - Magolha' ... ! Antes da debandada geral, Roberto fez (p. 286)
9	o guarda a dezer: "Haija quem mai' lance!" e um home brabado qu'aí havia, a picar ... a picar ... Os oitros (p. 310)
10	semp'e. Lá cuntaro até três e êl deu o nome à praça. O home par'cia rico: trajava à grave. Mãis aquilho era nau (p. 310)
11	Capitão chamou-me à ré e disse: "Pórigui! Falta-me um home à tub, charape! Vais tu ... ".' Dipois passei a um (p. 311)
12	vinte aqui. Inté demira qu'o cabelo nã se ponha a um home mais branco qu' aquela neve a rolar! - Lembro-me de (p. 312)
13	co lincinho dêle puinduirado na mão: «Toma lá, meu home! Andas sempre co êsse ranho a caiir. .. » «olha, olha, (p. 316)
14	... - Blós! Blós!! - Ah, Mariquinhas!, viste o meu home? ... Não havia dúvida: era o sinal de há um século, o bombão (p. 320)
15	Ai!, qu'os nossos nã chêgum a tempo! ... Qu'é do meu home? - Vossemecê está doida!? Então não sabe que o juiz já (p. 321)
16	qu' o alma do diabo de lia pulga, lia coisa qu'um home esmicha c'lia unha, havera de matar aquil senhor! ... (p. 376)

Tabela 5.7 - Contextos da palavra "home" no romance.

Completando esta inventariação, executamos novamente o *script* mas para a forma plural "homes" (*vide* Figura 5.6). Os resultados obtidos encontram-se descritos na Tabela 5.6. Notamos que as duas ocorrências da forma plural confirmam a predominância da pronúncia popular da palavra "homens" no romance.

```
>>>
Palavra a pesquisar: homes
Ignorar a capitalização? (S - 1 N - 0) 1
Número de caracteres à esquerda: 50
Número de caracteres à direita: 50
Nome que pretende atribuir ao ficheiro: homes50
Total 2 ocorrências.
>>>
```

Figura 5.6 - *Shell* da execução do `script5L` para a palavra "homes".

1	Criação Velha, arriba da pedra do vigia. Pra chomar os homes às canoas assoprava-se o búzio. 6 tempo! ... - Anda! (p. 131)
2	quer a sinhôra? ... Foi lia pegaça do pai ... Os homes são mum teimosos. Laiilêl dezia: «Se for lia piquena, (p. 376)

Tabela 5.8 - Contextos da forma plural "homes" no romance.

A pesquisa dos contextos de um dado vocábulo ou termo, que acabamos de ilustrar com o uso de "home/homes", resulta dos passos de programação elencados no **`script5L`**, ilustrado no Programa 5.8. Neste caso, a procura incidiu sobre 50 carateres antes da dita palavra e 50 depois, podendo o número de carateres escolhido variar largamente, conforme o tipo de texto em apreço e/ou o que se pretende com a pesquisa.

No âmbito do estudo das representações da natureza de ilhas, situadas em pleno Atlântico Norte, as referências ao "mar" e ao "oceano" assumem especial significado, ainda mais num livro que tem por título *Mau Tempo no Canal*.

Assim, a análise comparativa da expressão e significação destes dois vocábulos afigura-se como sendo, logo à partida, uma abordagem simultaneamente relevante e necessária.

Para prosseguir nesse sentido, a utilização do `script5L` leva-nos a descobrir os diversos enquadramentos das palavras "mar" e "oceano", agora pesquisados num raio de 100 carateres antes e depois de cada ocorrência. Como já foi observado, a escolha do número de carateres pode variar largamente, em consonância com o texto em estudo e/ou os resultados pretendidos. Assim, a linguagem de programação Python revela os diversos contextos e os significados possíveis dos sintagmas escolhidos, permitindo uma análise cuidadosa e sistemática moldada pelas potencialidades da tecnologia atual. Considerando os resultados desta pesquisa (*vide* Tabela 5.9).

1	História Insulana das Ilhas a Portugal Sujeitas no **oceano** ocidental. - Tu até sabes os títulos, Sophy! - Sofia sentou-se: (p. 129)
2	- Escuso impestores ... 1 Ingl. Artic ocean, **oceano** Glacial Árctico. 2 Ingl. Western Ground, Mar das Antilhas. (p. 287)
3	descampado! , nestas quatro paredes ... !») **oceano** Glacial Árctico (Em português do Pico: Ariôche) Àquela (p.309)

Tabela 5.9 - Contextos da palavra "oceano" no romance.

Constatamos que a palavra "oceano" figura apenas três vezes em todo o romance: uma como "oceano ocidental", no que é uma referência oblíqua ao Atlântico tirada do título de um livro de história "das Ilhas a Portugal Sujeitas", e duas vezes ligada, não ao Atlântico, mas ao oceano Árctico – "Em português do Pico: Arioche" – numa expressão tirada da história da baleação açoriana. Assim sendo, constatamos que o oceano Atlântico, matriz geográfica e geológica do arquipélago dos Açores, nem só fica diluído, como não figura como tal em parte alguma do romance.

Porém, segundo a perspetiva da ecocrítica, que privilegia o meio ambiente físico nas suas múltiplas manifestações e configurações, a ecologia terrestre depende visceralmente dos oceanos. Efetivamente, num planeta onde a água ocupa 70% da superfície terrestre e apenas 3% desta é água doce, as águas oceânicas são de tal forma preponderantes e essenciais que o microbiólogo Ed DeLong, por exemplo, foi levado a observar que o nome do planeta deveria ser "oceano" e não "Terra" (citado em Helmreich, 2009). No entanto, os resultados obtidos parecem indicar que o conceito de "oceano" está ausente da mundividência açoriana e que a própria palavra não faz parte do léxico normalmente utilizado nas ilhas, sendo substituída pelo conceito mais familiar e coloquial que brota da palavra "mar".

Na verdade, a palavra "mar" é recorrente no romance e foi pesquisada no contexto de cerca de 100 carateres à esquerda e à direita (*vide* Figura 5.7). Segundo os dados obtidos com o `script5L`, este vocábulo figura em 129 contextos, estando 40 registados na Tabela 5.10, e aponta para os aspetos da natureza açoriana que mais marcaram a ficção de Nemésio. Numa síntese possível, podemos constatar, por exemplo, que numerosas ocorrências conjugam o mar com a pedra vulcânica das ilhas, evocando a muito conhecida afirmação de Nemésio, descrevendo a açorianidade e o povo açoriano: "Somos de carne e de pedra. Os nossos ossos mergulham no mar" (Nemésio, 1932). Refletindo mais sobre o uso deste vocábulo, observamos que as ocorrências no final do romance, quando Margarida, recém-casada com André Barreto e de viagem para a Terceira, lança o seu enigmático anel de serpente ao mar, sublinham a força simbólica do ato da protagonista e a carga mitológica associada ao mar, patente noutra citação emblemática do autor: "As ilhas são o efémero e o contingente, só o mar é eterno e necessário" (Nemésio, 1928). Ainda no âmbito da Tabela 5.6, outros contextos do vocábulo "mar" (ocorrências 6, 8, 9, 11, 23, 27, 32 e 33) apontam para a importância do conjunto natural mar-ilha-em-frente e podendo levar-nos a pesquisar mais termos essenciais ao texto – "Canal" e "Pico", por ventura.

```
>>>
Palavra a pesquisar: mar
Ignorar a capitalização? (S - 1 N - 0) 1
Número de caracteres à esquerda: 100
Número de caracteres à direita: 100
Nome que pretende atribuir ao ficheiro: mar50
Ficheiro guardado com sucesso.
>>>
```

Figura 5.7 - *Shell* da execução do `script5L` para a palavra "mar".

1	curral da figueira. Tinham metido por ali, direitos à grota que bordava a quinta, da estrada até ao **mar**. O terreno ia ficando reduzido a pedra pura, com uns restos de vinha queimada da ressalga, figueiras-bravas, (p. 22)
2	parece um rapaz, sempre a pensar em regatas e em chás a bordo, com muitas raparigas. Mas gosto mais do **mar** que dos chás. Tu também? Ouviu-se o baque da maré, lançada com a certeza de mil homens que rolassem (p. 24)
3	abanavam com uma violência de mata abalada no centro. E, entre as árvores estaladas da quinta e o **mar** já grosso e tapado por uma pasta de escuridão, ficaram um bocado sufocados, sem poderem andar, voltados, (p. 25)
4	étape. Não estava um dia muito claro; mas vira aparecer o Sol dos lados da Terceira, todo sangrento num **mar** de chumbo, um mar como nunca tinha visto, fresco e sem nada que lhe cortasse a ilimitação parada, a (p. 36)
5	bordados: riam, riam disparatadamente e comiam peras com a casca. Tudo aquilo entrara por ela como o **mar** lentamente pela fuma, gorgolhando sem que se sentisse direcção ou violência - uma pura deriva de águas. (p. 37)
6	que estendia ali tanta coisa (quatro folhas!), e, sem a apaixonar, lhe dava a mesma impressão que o **mar** visto ao alto do Pico - uma vontade de o deixar subir sem saber para quê. Gostava muito de João Garcia? (p. 37)
7	costumava passar uma grande parte da tarde, dali seguia para os passeios discretos que dava na Rua do **Mar**, ao largo do Granel do avô. Margarida via o seu vulto vaga roso e magro perder-se para a boca da Doca, (p. 69)
8	Vinhas, quando o avô ainda se mexia e teimava em meter-se ao Canal. Em Fevereiro havia muitos dias de **mar** bravo, as lanchas afocinhavam nas grandes covas de água cavadas pelo vento da Guia. Para tirar o avô (p. 116)
9	restos do laranjal queimado da doença. Às vezes, na Primavera ou em certos dias de Outono em que o mar não deixava os barcos do Pico aproximarem-se da Horta e um vento puxado do Canal cobria o sobrado do (p. 150)
10	Feria sexta in parasceve Aquele dia santo no quartel pareceu a João Garcia uma ilha no **mar** da sua vida incolor, sem vento para vela alguma. Tocara à ordem mais cedo; a maior parte dos soldados (p. 157)

11	Naquele convento agora reduzido a quartel ao alto da cidade, defronte do Pico arroxeado das nuvens do **mar** ao entardecer, sentia-se na posse de uma força que vinha do fundo dos tempos, quando as ilhas não tinham (p. 160)
12	Angra com os sinos de meio arquipélago. Pela janela viam-se os calhaus rolados debaixo da toalha do **mar** prateado da foice da Lua. Ainda se pegavam sobre as divisões do saber e sua importância respectiva. (p. 199)
13	Capela e das saias rodadas das vizinhas da Rosa Bana. Sentia-se ali como a prancha que vem do alto **mar** e encontra enfim uma posição capaz de fixar as gaivotas e a sua própria massa de seivas, as suas fibras, (p. 212)
14	fichas de marfim da batota do seu principado microscópico. Ficavam ali secas, es tampadas, como enigmas do **mar** num cemitério. E nunca mais se esquecera daquela bicheza pequenina, Cucumaria abyssorum, encontrada pela (p. 218)
15	tão pouco irredutível nos seus velhos rancores, e aquela insistência do pai recomendando-lhe ares do **mar**, o sossego do campo, além da obra de caridade de dar ao tio comovido por aquele desenlace de drama familiar (p. 223)
16	sobre a estrada, com um caramanchãozinho de buganvílias que recebia de frente o bafo borrifado do **mar**. João Garcia sentara-se na banqueta de pedra, onde caía o rasto da noite azulada, excepcionalmente clara (p. 234)
17	acusando aquele toque: - Sim ... A administração da casa do comendador ajudou-me a aguentar no balanço. O **mar** levanta-se muito aí nas ilhas, Honório ... O nosso Canal é traiçoeiro ... Mas ... (vê lá tu o que são (p. 240)
18	Zimbreiro com a falésia brusca e cheia de bicos de basalto, que punha uma nódoa de tinta alastrada no **mar**. Uma tamargueira áspera da sua má vida ao vento marcava numa volta do caminho uma gruta de lava. Cheirava (p. 241)
19	bando de pombos bravos cobria o céu, e o seu levante soava esvoaçado e súbito na rocha aprumada ao **mar**. Uma fenda de lava atraiu a atenção de João Garcia. Estava mais contra a terra, abrigada pelo resto de (p. 246)
20	forrozinho de penas e de palhas vestia aquele recôncavo. E no meio da aspereza do Zimbreiro, entre o **mar** escuro da hora da tarde e da rocha empinada e negra, aqueles restos de ninhos pareciam um último apelo (p. 246)
21	pouco? .. Ouviu ... ? Outra vez. Parece senha. Deve ser agradável caçar assim ao anoitecer, com um **mar** que nem mexe ... O bote é airoso. Não se demore. Cuidado não caia ... - Então boa-noite ... - Boa-noite. (p. 248)

22	Vinhas, quando o avô ainda se mexia e teimava em meter-se ao Canal. Em Fevereiro havia muitos dias de **mar** bravo, as lanchas afocinhavam nas grandes covas de água cavadas pelo vento da Guia. Para tirar o avô (p. 263)
23	para uma menina. Mas o resultado da imprudência parecia luzir na casa, no dia claro que azulava o **mar** até à Horta e deixava negra e calma a pedra queimada do Pico - finalmente nos olhos baços de Manuel Bana (p. 264)
24	ter acudido a mê pai! - Por essa costa abaixo mete um respeito a todos! ... Morrê-le um filho no **mar** q' ando a baleia parida matou o baleote à rabada e revirou a canoa, e êl troixe a cumpanha a salvo noitro (p. 285)
25	bocado de tempo rofe ... Já passa. - Dês o oiça, meu amo! - disse o João da Cezilha. A gente veve do **mar** ... Se tamos munto tempo im terra, ía quarta de milho ainda s'amanha ... Tenho lá a inxada incostada à fruntal, (p. 288)
26	cerrado separava-se em pequenas massas soltas, progressivamente altas e móveis sobre a pedra queimada e o **mar** redondo. Pareciam aves fantásticas que o instinto reúne e convida a emigrar. Distinguiam-se as casas (p. 294)
27	misteriosa, que um resto de lua referia à cratera do Pico toda escorrida de lavas até ao nível do **mar**. Ouvia-se através dos corredores desertos o assobio de Pedro limpando a Manchester na cozinha. O Açor, (p. 307)
28	esbofando, nem podia dizer Deus te salve. A sua respiração acelerada e difícil balançava-lhe a barba como o **mar** levantado embala a embarcação; os seus olhos grandes e ingénuos bebiam o traço do horizonte. Entretanto (p. 325)
29	mestra? ... Além disso, o motor da lancha parecia agora pegado, e Roberto navegava um pouco a corta-**mar** em direcção à canoa onde descobrira Margarida. O vento avivara bastante; e a pouca vaga, rolando já larga (p. 326)
30	deixando-se levar. A canoa, às vezes adornada da manobra, roçava a borda no gume fresco e vivo do **mar**. As nuvens açorianas, a princípio paradas e aos pares, (Nuvens paradas, cor de cobre, E temporal que (p. 327)
31	enterrado no cachalote. Semelhante voz, ouvida naquela canoa comprada com dinheiro dos seus, naquele **mar** verde e belo que parecia o quintal da sua casa, solta por aquele homem agigantado e peludo e apesar (p. 328)
32	a cidade alarmada (um escândalo!)? ... dez ... vinte binóculos assestados das traseiras da Rua do **Mar** e dos altos da Horta sobre o Canal de São Jorge ... ? E ali sentada na pedra, com a face no punho esquerdo (p. 331)
33	flancos medonhos os futuros caminhos das chipeiras' e dos rnancins/. Um vago resto de dia convertera o **mar** num imenso pano de angrim, como se um gigante tivesse enfiado um par de alvarozes e se preparasse para (p. 332)

34	gente responde por tuido ...' O Intavante chegou com uma grande baila' de água salgada na mão: - O **mar** tá rofe, ti Amaro ... lia carneirada solta. - De pantalonas arreadas até acima dos joelhos, as solas (p. 335)
35	grande lago de sangue que enodoava a água e ia invadindo tudo: mãos, homens, canoa, o monstro e o **mar**, e que lhe entrava pelos olhos dentro como se visse de repente golfar o coração de tudo o que tinha amado (p. 342)
36	Mas, sentindo arfar docemente a barba do ti Amaro de Mirateca, cheia do hálito vivo que entrava do **mar** pela fuma dentro, ganhou confiança e adormeceu. Xácara dos mortos-vivos Não, Sr.ª Baronesa; esta (p. 343)
37	Noiva de quem?! ... Ah! Agora, sim; agora é que a resposta estava talvez dada para sempre, riscada no **mar** pela proa de uma canoa do Pico e pela escolha cega de uma baleia trancada ... Noiva de André Barreto! (p. 356)
38	passasse um ferro de engomar. Pelas fendas abertas no chão percebia-se um urro longínquo, troado, que o **mar**, de crista rente às falésias, vinha tornar mais próximo. E a lava ardente, pastosa, gorgolava devagar (p. 359)
39	Mateus Dulmo. Graças a estes cálculos tão fáceis ao alto da Urzelina, no silêncio húmido do campo e do **mar** barrado de ilhas, Margarida podia sentir-se livre de realizar a única aspiração que verdadeiramente ganhara (p. 362)
40	caracolavam a íngreme subida das falésias que deixavam a Fajã das Almas triste e enrascada no **mar**, ela já não seria a Bidinha filha do genro do Sr. Charles Clark, arribada a São Jorge numa baleeira fugi (p. 393)

Tabela 5.10 - Contextos de 40 das 129 ocorrências da palavra "mar"

Ambos nomes próprios escritos com letra maiúscula, o Canal e o Pico protagonizam, em conjunto, o espetáculo paisagístico da natureza açoriana que se desenrola, a toda a hora, perante o mirar privilegiado da cidade da Horta, "*um camarote de frente para aquele palco de todo o ano*" (Nemésio, 2001: 34). Ao referir esta paisagem única mais à frente no romance, a voz narrativa destaca a carga simbólica e mítica da natureza dos Açores, força anímica e autónoma, anterior a qualquer presença humana, quando observa que (ocorrência 1 da Tabela 5.10) "a cidade, defronte do Pico arroxeado das nuvens do mar ao entardecer, sentia-se na posse de uma força que vinha do fundo dos tempos, quando as ilhas não tinham ainda sinal de nada humano" (p. 160).

Neste âmbito, a Tabela 5.11 elenca os contextos, num raio aproximado de 100 carateres à esquerda e à direita, de algumas das ocorrências da palavra "Canal", usada 65 vezes, e do vocábulo "Pico", presente em 137 "ocasiões ao longo do romance. Estas ocorrências podem ser consultadas separadamente ou em conjunto, coincidindo por vezes com o uso do vocábulo "mar" já referido. Aproximando-nos do fim desta reflexão sobre a utilização da linguagem Python na análise do romance *Mau Tempo no Canal*, constatamos que, entre os resultados obtidos, figura o sintagma que deu o título a este grande romance (ocorrência 4 da Tabela 5.11), também ele "concebido como matéria de um mito, ligado à rocha das ilhas e às noites de mau tempo no Canal" (p. 177).

1	esboroada, marcando o sentido do enxurro. E, para lá do «calhau» coberto das mantas da cheia, via-se o **Canal** ainda amargo, com o Pico negro e cónico ao fundo. O cano do Funchal fumava a meio da Doca, entre um (p. 31)
2	que os passara na Pedra da Burra, nas Vinhas, quando o avô ainda se mexia e teimava em meter-se ao **Canal**. Em Fevereiro havia muitos dias de mar bravo, as lanchas afocinhavam nas grandes covas de água cavadas (p. 115)
3	de outono em que o mar não deixava os barcos do Pico aproximarem-se da Horta e um vento puxado do **Canal** cobria o sobrado do cartório de grave tos e folhas secas, Januário mandava sair toda a gente, Ângelo (p. 150)
4	querer-bem concebido como matéria de um mito, ligado à rocha das ilhas e às noites de mau tempo no **Canal**. Assim perdera o segredo daquela ocasião de uma valsa, como quem deixa cair uma minhoca inevitável debaixo (p. 177)
5	no Pico, lá para Santo Amaro. De vela latina e giga, aparelhada à baleeira, a lancha bolinou no **Canal**. Margarida, há meia hora de binóculo no torreão da Poça, viu-a descrever um largo S ao vento, de borda (p. 214)
6	comendador ajudou-me a aguentar no balanço. O mar levanta-se muito aí nas ilhas, Honório ... O nosso **Canal** é traiçoeiro ... Mas ... (vê lá tu o que são as coisas do mundo!) Todos dizem que estou (p. 240)
7	ser pescadores. Margarida, com a cabeça nas mãos apoiadas aos vidros, espiava a mínima pulsação do **Canal** barrado de negrume; as luzes da Horta e o céu largamente estrelado reflectiam-se nas águas volumosas (p. 251)
8	disse Margarida como de si para si, estremecendo. - Constipou-se talvez na lancha ... ! a aragem do **Canal** é traiçoeira. Tem havido tanto caso de peste no Faial. .. ! Não nos bastava o susto que apanhámos com (p. 377)

Tabela 5.11 - Contextos de algumas das ocorrências da palavra "Canal".

Desta forma, tentamos exemplificar e explicar algumas das potencialidades da mineração de texto no contexto da análise linguística e literária, nomeadamente no estudo do romance *Mau Tempo no Canal* de Vitorino Nemésio. Esperamos que as situações e explicações apresentadas sejam úteis para os colegas a trabalhar em Humanidades. Em abono da verdade, as possibilidades são muitas e dependem da criatividade e iniciativa do investigador que aceite o desafio da integração das novas tecnologias na sua atividade.

Bibliografia

Brandão da Luz, J. (2009). "Percursos da mundividência naturalista nos Açores" in *Reflexão sobre Mundividências da Açorianidade*, (Coord.) Castro, G., Miúdo, B. & Carvalho, M. (pp. 13-34). Ponta Delgada: Universidade dos Açores.

Carvalho, M. (2006). *A natureza em Antero de Quental: O projecto de uma metafísica positivista.* Lisboa: Imprensa Nacional-Casa da Moeda.

Cruz, J. V. (2009). *Crónicas do ambiente.* Mirandela: João Azevedo Editor.

Fagundes, F. (1998). "*Traduzindo Mau Tempo no Canal para o inglês*" in *Vitorino Nemésio: Vinte anos depois,* (Coord.) Machado-Pires, A.,

Garcia, J. M. (1994). *Introdução, Mau Tempo no Canal, Vol. III, Vitorino Nemésio: Obras Completas* (pp. 7-15). Lisboa: Imprensa Nacional-Casa da Moeda.

Helmreich, S. (2009). *Alien Ocean: Anthropological Voyages in Microbial Seas.* Berkeley: University of California Press.

Lepecki, M. L. (1974). "*Sobre Mau Tempo no Canal" in Críticas sobre Vitorino Nemésio,* (pp. 167-179). Lisboa: Bertrand.

Mendes, V. (2009). "*A ecocrítica e os Açores" in Reflexão sobre Mundividências da Açorianidade*, (Coord.) Castro, G., Miúdo, B. & Carvalho, M. (pp. 149-157). Ponta Delgada: Universidade dos Açores.

Mourão-Ferreira, D. (1974). "*Sobre a obra de Vitorino Nemésio"* in *Críticas sobre Vitorino Nemésio*, (pp. 1-7). Lisboa: Bertrand.

Nemésio, V. (2001). *Mau Tempo no Canal* (1ª edição 1944). Lisboa: Leya, S. A. Sociedade Portuguesa de Autores e Publicações Dom Quixote.

Nemésio, V. (1932). "*Açorianidade*" in Ínsula, nº 7: Julho-Agosto.

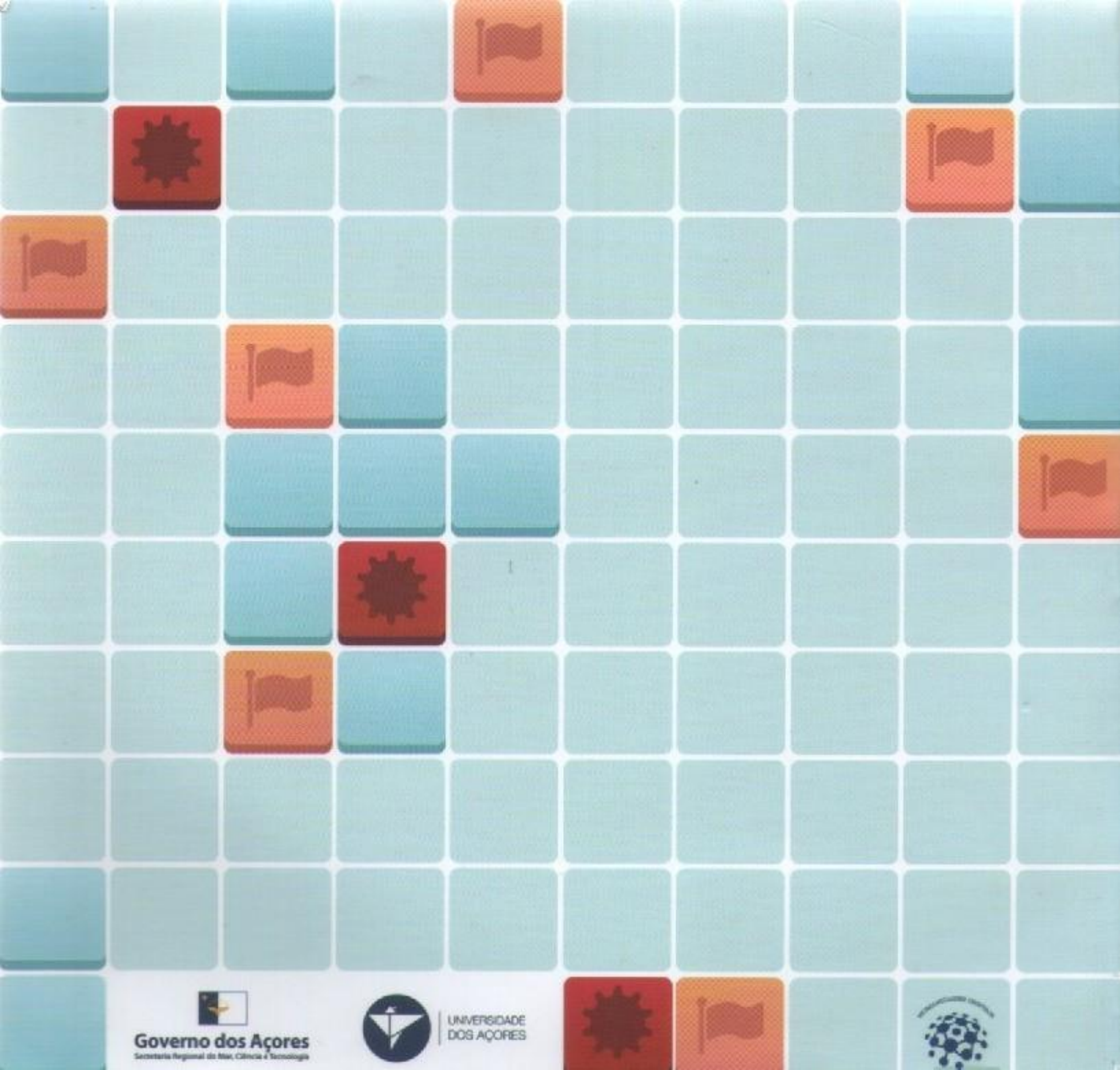
Nemésio, V. (1928). *O açoriano e o Açores.* Porto: Renascença Portuguesa.

Silva, H. (1985). *Açorianidade na Prosa de Vitorino Nemésio: Realidade, Poesia e Mito*. Lisboa: Imprensa Nacional.

Simas, R. (2014). "Frances Dabney e Samuel Longfellow: A natureza dos Açores numa perspetiva feminista e ecocrítica" in Simas, R. (Coord), *A Vez e A Vez e a Voz da Mulher: Relações e Migrações* (pp. 83-96). Lisboa: Edições Colibri.

Simas, R. (2013). "Cartas e desenhos evocando a natureza dos Açores: Um olhar ecocrítico sobre Samuel Longfellow na Horta" in Costa, R. M. (Coord), *Boletim do Núcleo Cultural* da Horta, Volume 9: 105-121.

Simões, J. G. (1974). "Mau Tempo no Canal" in *Críticas sobre Vitorino Nemésio,* (pp. 70-77). Lisboa: Bertrand.


Governo dos Açores
Secretaria Regional do Mar, Ciência e Tecnologia
UNIVERSIDADE
DOS AÇORES