

Os Arabismos próprios da língua portuguesa em época contemporânea: análise do uso comunicativo dos itens lexicais através das ocorrências autênticas em corpora eletrónicos

The Usual Arabisms in Portuguese Today: Analysis of the Communicative Use of Lexical Items through Authentic Occurrences in Electronic Corpora

MARIA ANTONIETTA ROSSI [rossi.mariaantonietta@unistrasi.it]
Università per Stranieri di Siena, Itália

RESUMO

Os arabismos representam uma parte importante do património histórico lusófono, uma vez que muitos itens lexicais entraram a partir do século VIII – período que atesta o início do duplice processo de islamização e arabização da Península Ibérica – para denotar todas as inovações culturais e sociais introduzidas pelos novos dominadores. A partir do levantamento lexicográfico por nós efetuado em 2011, o presente estudo visa verificar quais arabismos peculiares do idioma português são ainda utilizados nas práticas comunicativas contemporâneas através da análise tanto qualitativa como quantitativa das ocorrências de cada lema nos textos do corpus eletrónico *Now* (*News on the Web*).

PALAVRAS-CHAVE

Arabismos; Competência Comunicativa; Corpora Eletrónicos; Lexicografia; Semântica Descritiva

ABSTRACT

Arabisms represent an important part of the Portuguese-speaking historical heritage, since many lexical items entered in the 8th century - a period that attests the beginning of the double process of Islamization and Arabization of the Iberian Peninsula - to denote all the cultural and social innovations introduced by new dominators. Based on the lexicographic survey carried out in 2011, the present study aims to verify which peculiar arabisms of the Portuguese language are still used in contemporary communicative practices through both qualitative and quantitative analysis of the occurrences of each lemma in the texts of the electronic corpus *Now* (*News on the Web*).

KEYWORDS

Arabisms; Communicative Competence; Electronic Corpora; Lexicography; Descriptive Semantics

RECEBIDO 2021-03-29; ACEITE 2021-05-18



1. Introdução: a componente árabe do léxico português

Os vestígios arábicos na língua portuguesa constituem uma parte fundamental do património histórico e cultural do espaço lusófono, uma vez que os itens lexicais baseados em etimologias ligadas às variações diatópicas do mundo islâmico entraram no uso comunicativo dos povos ibéricos a partir provavelmente do início do século VIII: nesta altura teve início, de facto, o duplice processo de islamização e de arabização dos territórios meridionais de Espanha e Portugal, rebatizados respetivamente com os topónimos *Al-Andalus*, a atual Andaluzia, e *Gharb al-Andalus*, correspondente à região do Algarve (Alves 1999: 56; Picard 2000: 7). Nestas áreas desenvolveu-se, até ao século XIII, “a brilhante civilização árabo-muçulmana” (Alves 1999: 48), onde o emprego conversacional do novo idioma era sinónimo de prestígio cultural (Suisse 2020: 47), atitude esta que decretou, com efeito, a entrada e o emprego de numerosos arabismos, apesar de já existirem, na bagagem lexical das comunidades ibéricas, formas românicas anteriores com a mesma denotação sémica, como argumenta Maria José de Moura Santos (1980: 584).

O fenómeno do empréstimo linguístico dos falares árabes, oriundos nomeadamente das áreas setentrionais da África – mais próximas a nível geográfico aos territórios do Sul da Península Ibérica e, por conseguinte, facilmente acessíveis pelos grupos berberes – incentivou, durante cinco séculos, o processo de dinamismo lexical das línguas neolatinas, neste período em fase de evolução morfossintática e fonética por causa da miscigenação entre o latim vulgar importado pelos Romanos e as línguas tanto de substrato como de adstrato e superstrato (Castro 1991: 139–159; Maranhão 2018: 131). Para denotar todas as inovações culturais introduzidas pelos povos muçulmanos, nomeadamente tecnicismos pertencentes ao campo semântico da arte bélica islâmica, da estrutura político-social oriental, das ciências e do sector agrícola, foram adotadas inéditas etiquetas lexicais que constituíram progressivamente parte da herança linguística oral das comunidades ibéricas (Santos 1980: 578; Alves 1999: 5–8; Machado Filho 2013; Suisse 2020).

A interiorização dos novos lemas através da oralidade foi incentivada nomeadamente pelos falantes moçárabes (Santos 1980: 578; Picard 1983; Mattoso 1985; Samu 2010: 49; Maranhão 2018: 134; Suisse 2020: 47), “*actifs acteurs sociaux*” (Picard 2000: 98) que dominavam fluentemente a *al-garavia*, mas que assimilaram as unidades lexicais aglutinando o artigo definido de género masculino *al*, não sabendo discernir as duas entidades, ativando, desta feita, adaptações fonológicas às línguas neolatinas de chegada (Samu 2010: 50; Suisse 2020: 56).

Tal convivência pluricultural e multilíngue (Guichard 1977; Maranhão 2018) perdurou, em território lusófono, até 1275 com o fim do processo de Reconquista (Picard 2000: 107–129), mas se manifestou novamente no Oriente durante os séculos dos Descobrimentos, época que, para o estudioso Suisse (2020: 48), corresponde à terceira fase histórica de entrada de arabismos na língua portuguesa, uma vez que o contacto diuturno entre o Português Moderno e as línguas “exóticas” faladas pelos povos africanos e asiáticos conquistados (Santos 1980: 579; Buescu 1983) gerou, mais uma vez, o fenómeno do empréstimo linguístico do árabe.

Este processo secular de importação lexical – e de consequente evolução semântica – implicou a introdução, conforme o levantamento lexicográfico por nós publicado em 2011 (Rossi 2011: 147–163), de 1.009 arabismos, dos quais mais da metade apresenta um lema equivalente em espanhol, cujo léxico de origem árabe compreende, pelo contrário, quase 4.000 entradas (Alves, 2013: 23).

Tendo em consideração este estudo lexicográfico, o objetivo principal desta investigação consiste em ampliar a análise realizada em 2011, verificando, por conseguinte, o real emprego comunicativo dos arabismos peculiares do idioma lusitano em época contemporânea, através do exame das ocorrências das unidades lexicais selecionadas em textos autênticos digitais recolhidos em corpora eletrónicos. Graças a esta tipologia de investigação, o presente trabalho visa identificar, por um lado, os itens de origem árabe atualmente ativos em perspetiva socio-pragmática no português hodierno e, por outro, explorar os lexemas que, de modo inverso, caíram em desuso, avançando hipóteses quanto às razões pelas quais estes arcaísmos foram substituídos no uso comunicativo, em perspetiva diacrónica, por vocábulos de etimologia latina.

2. Metodologia de pesquisa e embasamento teórico

As entradas que embasam o presente estudo – 267 arabismos próprios do idioma lusitano que não foram atestados em castelhano – foram identificadas através do levantamento por nós efetuado em 2011 (Rossi 2011) baseado na colação de 7 dicionários etimológicos da língua portuguesa e espanhola¹, processo que permitiu, como já antecipámos na parte introdutória, recolher 1.009 unidades lexicais de nosso interesse. O objetivo principal deste inventário foi catalogar, nesta primeira etapa da investigação, os arabismos peculiares do português, verificando se cada entrada apresentava um lema equivalente em espanhol através do auxílio do dicionário monolíngue online da RAE (Real Academia Española)².

Esta primeira análise apresentou, conseqüentemente, um panorama linguístico diacrónico dos lemas baseado no exame do século de atestação de cada arabismo, mas não foi verificado o respetivo uso no domínio comunicativo do português contemporâneo, objetivo primário, para tanto, desta segunda fase da pesquisa.

De facto, os autores por nós consultados (Souza 1830; Vasconcelos 1956; Machado 1958, 1956–59; 1991; Corriente 2003; Alves 2013; Suisse 2020: 49) focalizaram-se nomeadamente na recolha das entradas e na reconstrução da exata etimologia destas unidades verbais, mas não completaram esta tipologia de trabalho identificando quais arabismos ainda estão ativos na dimensão lexical do português hodierno e quais, ao invés, caíram em desuso, tornando-se, desta feita, em arcaísmos. Com efeito, a mais recente e prestigiosa obra lexicográfica neste âmbito, i.e. o *Dicionário de Arabismos da Língua Portuguesa* publicado em 2013 por José Adalberto Coelho Alves, por nós consultada para atualizar o levantamento de 2011, se por um lado apresenta um copioso número de entradas – 18.073 (Alves 2013: 23) – que incluem vocábulos, antropónimos e topónimos reconstruídos meticolosamente sob o ponto de vista etimológico (enquanto nas obras anteriores

1 Os 7 dicionários etimológicos e monolíngues do corpus são os seguintes, em ordem cronológica de publicação: 1) Sousa, J. D. (1830). *Vestígios da língua arabica em Portugal*. Lisboa: Academia Real das Sciencias; 2) Nascentes, A. (1955). *Dicionário Etimológico da Língua Portuguesa*. Rio de Janeiro: Livraria Académica; 3) Machado, J. P. (1956–1959). *Dicionário etimológico da língua portuguesa*. Lisboa: Editorial Confluência. 2 voll.; 4) Viterbo, J. D. S. R. (1965–1966). *Elucidário das palavras, termos e frases que em Portugal antigamente se usaram*. Porto/Lisboa: Civilização. 2 voll.; 5) Corominas, J. (1984–1991). *Diccionario crítico etimológico castellano e hispánico*. Madrid: Gredos. 6 voll.; 6) Machado, J. P. (1991). *Vocabulário Português de Origem Árabe*. Lisboa: Editorial Notícias; 7) Houaiss, A. (2001). *Dicionário Eletrónico da Língua Portuguesa*. Rio de Janeiro: Editora Objetiva.

2 <https://dle.rae.es/>

resultava ainda “obscura” a origem da maior parte das entradas), por outro lado, aliás, não explícita a lista dos arabismos ainda em uso na prática conversacional, nem a data de atestação de cada lema, fundamental para individuar a época histórica durante a qual o arabismo entrou no espaço linguístico lusófono.

Levantamento	Número dos lemas	Arabismos peculiares do português	Arabismos em comum entre português e espanhol
2011	1.009	267	742
2013	18.073*	Não identificados	Não identificados

* O amplo número de entradas presente no trabalho de Alves (2013) deve-se ao registo quer de itens lexicais, cujas numerosas variantes ortográficas são lematizadas em separado, quer de antropónimos e topónimos, campo semântico, este último, que foi particularmente profícuo em termos de criação verbal durante o período de dominação islâmica da Península Ibérica. Merece a pena destacar que o presente dicionário não descreve nem o corpus de obras lexicográficas consultadas para elaborar o levantamento, nem o ano ou o século de registo de cada entrada.

Figura 1. Tabela contrastiva entre os levantamentos de 2011 e de 2013

Para completar a análise, por conseguinte, baseamo-nos também no *Dicionário etimológico da língua portuguesa* de José Pedro Machado (1956–59), uma vez que é o único instrumento lexicográfico a apresentar a data e o século de atestação, com a respetiva fonte literária e documental de cada entrada, dados não presentes, aliás, na obra de Alves (2013). Informações estas indispensáveis para compreender, eventualmente, a razão pela qual a maior parte dos arabismos foi substituída diacronicamente por unidades lexicais de origem latina, que já existiam antes da chegada dos povos muçulmanos.

Em decorrência disto, a fim de completar o estado da arte discutido, é do nosso principal interesse avançar nesta segunda fase da investigação lexicográfica apurando o emprego comunicativo concreto dos 267 arabismos peculiares do português: para levarmos a cabo este estudo de caso sincrónico, recorreremos ao corpus *NOW (News on the Web)*, disponível online³, escolhido por reunir diferentes géneros digitais redigidos no ciberespaço em língua portuguesa entre 2012 e 2019, nomeadamente artigos e notícias que pertencem – considerando a classificação textual proposta por Werlich (1975) – à tipologia argumentativa e informativa, categorias em que mais se manifestam as mudanças renovadoras da língua, nomeadamente no âmbito semântico e lexical.

Através duma abordagem analítica de tipo quer i) qualitativo (Lüdke, André 1986) – finalizada a identificar os campos semânticos de referência dos arabismos peculiares mais utilizados de molde a avançar hipóteses quanto à resistência comunicativa destas entradas no léxico fundamental da língua portuguesa – quer ii) quantitativo (Biderman 1978; Turato 2004) – cujo objetivo consiste em elaborar dados numéricos e estatísticos em relação ao tema abordado –, foi possível distribuir os arabismos peculiares, através do exame do número de ocorrências de cada entrada nos textos autênticos do corpus eletrónico escolhido, em 4 faixas – por nós numericamente ela-

3 <https://www.corpusdoportugues.org/now/>

boradas⁴ – conforme a respetiva frequência de uso: 1) Alta (mais de 800 ocorrências), 2) Média (de 700 a 100 ocorrências), 3) Baixa (de 100 a 20 ocorrências) e 4) Desuso (de 20 a 0 ocorrências).

Com base nos resultados conseguidos, apresentados no parágrafo a seguir, identificaremos i) os lemas pertencentes às quatro categorias supracitadas, ii) os arabismos peculiares que caíram em desuso e iii) as possíveis razões, sob o ponto de vista tanto linguístico como histórico, que determinaram o desaparecimento na linguagem corrente destes arcaísmos e a respetiva substituição por verbetes de origem latina, apresentando um exame contrastivo entre as datas de atestação fornecidas pelo dicionário etimológico de José Pedro Machado (1956–59), não existindo por enquanto, a este respeito, outra obra lexicográfica completa da língua portuguesa.

3. Os arabismos próprios da língua portuguesa: itens lexicais ativos e inativos em época contemporânea

A análise exploratória por nós conduzida, para verificar o real emprego comunicativo dos 267 arabismos, forneceu resultados interessantes quanto à evolução diacrónica dos itens lexicais postos em destaque nesta investigação a nível tanto semântico como pragmático-conversacional.

Considerando o número de ocorrências de cada etiqueta lexical nos textos autênticos digitais do corpus eletrónico *NOW*, podemos asseverar que, atualmente, o número dos termos de origem arábica ainda em uso no desempenho comunicativo, tanto do Português Europeu (PE) como do Português Brasileiro (PB), é bastante reduzido, uma vez que apenas 39 entradas resultam ativas, i.e. 15% do total dos arabismos peculiares⁵. Dentro deste limitado grupo lexical, podemos fazer uma distinção entre a) arabismos de alta, média e escassa frequência conforme o índice de ocorrências obtido e b) arcaísmos por um total de 228 (85%) – que caíram em desuso, uma vez que não recorrem em nenhum texto do corpus consultado, mas que fazem parte, pelas razões históricas anteriormente apresentadas, do património linguístico do espaço lusófono: estas primeiras informações quantitativas permitem, então, aventar a hipótese de que a maior parte dos arabismos peculiares resulta inativa no estado atual da lexicologia portuguesa.

Os 10 verbetes pertencentes à categoria de alta frequência (de 300.000 a 800 ocorrências) resultam ser, em ordem decrescente, *fato* (239.097)⁶ – o arabismo mais utilizado no português hodierno –, *fatia* (15.069), *algarvio* (4.663), *alface* (3.089), *sucata* (2.559), *algoz* (2.368), *alicerce* (1.365), *giz* (1.348), *cabide* (831) e *azáfama* (801). Em seguida, identificámos 13 arabismos de média frequência (de 700 a 100 ocorrências) – *arraia* (469), *alcatra* (454), *açorda* (439), *algema* (415), *alcateia* (238), *saloio* (228), *aldrabão* (212), *regueifa* (210), *algibeira* (171), *alguidar* (170), *ceifa* (161), *acém* (144) e *alqueire* (108) – e 16 termos de baixa frequência (de 90 a 10 ocorrências) – *ceroulas* (80), *alforreca* (69), *adiafa* (68), *gazua* (55), *haje* (52), *alcofa* (47), *aluá* (44), *anafado* (38), *enxovia* (37),

4 Não existindo um modelo numérico universal de referência, a proposta classificatória apresentada foi elaborada tendo em consideração o trabalho de Maria Tereza Camargo Biderman, baseado na análise do léxico de frequência do Português Brasileiro (Biderman 1998).

5 Na análise apresentada, não se fez uma distinção das ocorrências entre o PE e o PB.

6 Neste total, expurgámos as ocorrências com o significado de “facto” para o PB.

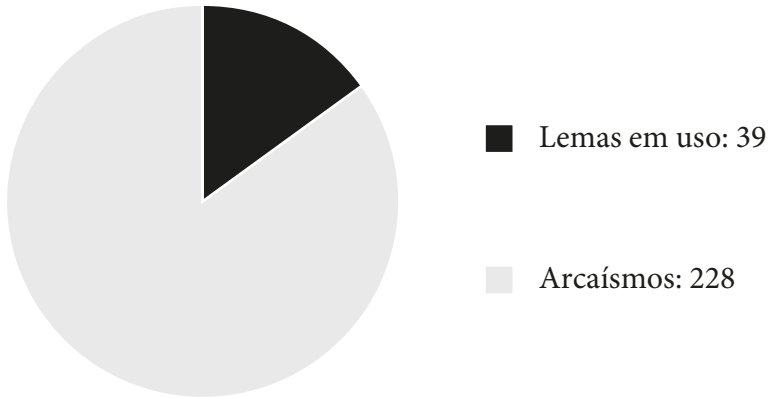


Figura 2. Percentagem dos arabismos em uso e em desuso

mafarrico (32), *frangão* (30), *taibo* (27), *arraia-miúda* (23), *aldebarã* (21), *osga* (21) e *azurracha* (0) – que completam este quadro analítico quantitativo⁷.

Vale especificar que, dentro desta última categoria lexical, foi igualmente incluído, apesar de não aparecer em nenhum dos textos dos corpus, o lema *Azurracha* – correspondente a uma tipologia de embarcação empregada para navegar nas águas do rio Douro – uma vez que, como referem os únicos dicionários da língua portuguesa que reportam a entrada – *Infopédia* e *Priberam*⁸ – são unidades lexicais ativas unicamente nas variações diastráticas da área geográfica setentrional de Portugal, onde este arabismo representa, no nosso parecer, um tecnicismo próprio da tradicional arte piscatória desta área. Quanto a este termo – cujo emprego conversacional resulta circunscrito, segundo as informações encontradas, nas comunidades de pescadores do Norte – não é possível avançar hipóteses quanto ao período de entrada nas terras peninsulares, uma vez que nenhum dicionário etimológico consultado apresenta uma data de provável introdução do lema na língua portuguesa.

Entre os arabismos próprios, ainda ativos no português contemporâneo, encontramos lemas antigos que entraram no uso comunicativo dos povos ibéricos durante a fase de dominação muçulmana que, segundo Abdelilah Suisse (2020: 47), corresponde ao segundo período de entrada destas etiquetas inovadoras na Península, i. e. dos finais do século VIII ao século XIII: esta época, de facto, resulta mais dinâmica em relação ao fenómeno de importação lexical, uma vez que os falantes locais interiorizaram a maior parte dos arabismos necessários para denotar todas as inovações socioculturais introduzidas com o processo de islamização dos territórios ocupados, fenó-

⁷ Tendo em consideração os dados identificados, alguns arabismos, utilizados nomeadamente em contextos socio-discursivos específicos e setoriais (como, a título de ilustração, *algeroz* e *osga*), resultam ter uma baixa frequência porque não recorrem com assiduidade na língua corrente.

⁸ *Infopédia* e *Priberam* são atualmente os principais dicionários online da língua portuguesa de domínio público (<https://www.infopedia.pt/>; <https://dicionario.priberam.org/>).

meno que tornou bilingues os falantes ibéricos (Suisse 2020: 47). Com efeito, nesta fase de evolução linguística dos falares neolatinos foram registados também *Alqueire* (1111), *Teiga* (1175), *Regueifa* (1177) e *Alcofa* (1188), arabismos arcaicos que, apesar de terem entrado na prática conversacional no século XII, ainda resultam ativos no português hodierno conforme os resultados obtidos através da análise quantitativa das ocorrências nos textos do corpus NOW.

O lema *alqueire*, não obstante se refira a uma antiga unidade de medida para secos e líquidos, cujo uso estava circunscrito sobretudo ao Sul da Península (Ribeiro 1965: 180–187; SANTOS 1980: 580), ainda resulta ativo, sendo utilizado sobretudo em âmbito regional no Português Brasileiro para indicar uma medida agrária correspondente, segundo o dicionário online Priberam, a 48.400 metros quadrados. Como informa Machado (1956–59, s. v. *alqueire*), esta entrada foi registada em 1111, num documento administrativo respeitante à gerência da área coimbrã, que se encontra na parte *Leges et Consuetudines* da coletânea *Portugaliae monumenta histórica* (Herculano 1863: 356), onde se menciona o *alqueire* como medida utilizada, de facto, para os alimentos. Ao longo dos séculos, porém, esta entrada passou por um processo de evolução semântica adquirindo uma nova aceção denotativa, já que o dicionário etimológico de Antônio Geraldo da Cunha (2011: s. v. *alqueire*), pertencente à lexicografia de matriz brasileira, refere ser uma medida quer de capacidade, quer de área.

Quanto ao antigo termo *regueifa*, pertencente ao campo semântico da culinária – que os povos islâmicos enriqueceram imensamente durante os cinco séculos de dominação – e referente a uma tipologia de pão em forma de rosca realizado com farinha de trigo, é um termo que recorre frequentemente, conforme os resultados obtidos neste levantamento, em receitas pertencentes à categoria dos textos injuntivos. É um verbete que sempre pertenceu a este âmbito lexical, como testemunha o documento em que foi atestado pela primeira vez, mencionado por Machado (1956–59: s. v. *regueifa*) no seu dicionário, i. e. *O título de doação de herdades que fez Lourenço Gomes a sua mulher Goncina Paes*, manuscrito datado 22 de novembro de 1177, onde o homem elenca uma lista de objetos que tenciona deixar como herança à sua esposa, que inclui vinho e 5 *rejeifas*⁹.

O campo semântico da culinária é, com efeito, um dos setores em que os arabismos resultam particularmente ativos, o que é perceptível na lista acima apresentada relativamente às entradas ainda hoje em uso: a maior parte foi atestada na língua portuguesa, considerando as informações apresentadas nos dicionários examinados (Machado 1956–59; Houaiss 2001; Cunha 2011), no século XVI – tal como *acém*, *açorda*, *adiafa*, *alcatra*, *alface* e *fatia* – período que, como argumenta Abdelilah Suisse (2020: 47), constitui a terceira fase de entrada dos arabismos, quando, através das campanhas de exploração territorial do Oriente promovidas pela dinastia dos Aviz (1385–1580), gera-se novamente o contato de interferência linguística entre o português e os falares arábicos do continente asiático e africano. A atestação de *aluá*, pelo contrário, resulta ainda duvidosa, uma vez que Machado (1956–59: s. v. *aluá*) refere que o verbete foi anotado no século XVI, mas Houaiss (2001: s. v. *aluá*), porém, informa que foi registado em 1727.

Os lemas de media frequência *acém* (a parte anterior da lombada dos bovinos), *açorda* (prato a base de pão, temperado com azeite, alho e ervas aromáticas) e *alcatra* (carne bovina extraída da

9 Cfr. Anónimo. (1991). O título de doação de herdades que fez Lourenço Gomes a sua mulher Goncina Paes. In *Revista Lusitana*. Lisboa: Livraria Clássica Editora, vol. XIV, p. 256.

parte posterior do dorso), recorrentes nomeadamente em receitas culinárias, resultam igualmente ativos no léxico gastronómico lusitano, dado que não existem correspondentes unidades sinonímicas de origem latina na prática discursiva hodierna: isto faz pressupor que, antes da chegada dos povos islâmicos nas terras ibéricas, não se utilizasse esta tipologia de comida e, para tanto, foi necessário adquirir estes itens lexicais para conotar os novos hábitos alimentares, adotados sobretudo pelas comunidades moçárabes. Em relação ao lema *adiafa*, regionalismo que é utilizado para indicar a refeição oferecida aos trabalhadores depois do fim da vindima, podemos afirmar que é um arabismo utilizado apenas nas variações diatópicas e diastráticas dos territórios do Alentejo e do Algarve, como referem os dicionários *Infopédia* e *Priberam* (s. v. *adiafa*): em decorrência disto, o emprego comunicativo deste arabismo é circunscrito aos contextos interativos rurais da zona meridional de Portugal, circunstância que determina, por conseguinte, a escassa utilização comunicativa deste termo, tal como *azurracha*. Também os verbetes *aluá* (massa de açúcar e mel utilizada para a preparação de doces) e *frângão* (galo jovem) são entradas com baixa frequência de uso, razão pela qual estão a desaparecer gradualmente da bagagem lexical contemporânea, circunstância que já se verificou, aliás, com *alfêloa* (sinónimo de *aluá*) e outras etiquetas verbais – assim como *adâçama* (gordura dos alimentos), *adora* (tipologia de milho cultivada no Magrebe), *alçuz* (sinónimo de *alçaçuz*), *alfaco* (tipologia de cogumelo de chapéu vermelho), *alfava* (variante ortográfica do arabismo *alfavaca*), *algaz* (equivalente a *tâmara*, arabismo atestado anteriormente, no século XV), *anafa* (que conviveu com o termo de origem latina *meliloto*), *anáfega* (sinónimo de *açoifeifa*), *atequipera* (variedade de pêra que, hoje em dia, se designa por *fidalga*), *azevizinhos* (espécie de prato a base de cuscuz), *bácoro* (equivalente a *porquinho*), *batafaluga* (a semente do anis), *bateca* (sinónimo de *melão*, este último utilizado em área lusófona a partir do século XI¹⁰), *buza* (uma espécie de bebida fermentada com cereais consumida sobretudo no Oriente), *farroupo* (outro arabismo que designa o *porquinho*), *fístico* (o fruto da pistácia), *maçal* (o soro de leite) e *xarifa* (sinónimo de *figo*¹¹) – arabismos próprios da língua portuguesa que se tornaram, com efeito, em arcaísmos, por terem sido utilizados apenas em circunstâncias históricas culturalmente circunscritas, sem entrar no real uso comunicativo, desaparecendo gradualmente no espaço lusófono com o fim da dominação islâmica e do período secular de bilinguismo.

No âmbito da culinária, o lema mais utilizado no português contemporâneo resulta ser *alface*, dado que a cifra de ocorrências relevada corresponde a 3.089. Não obstante o cultivo desta hortaliza, típico da região mediterrânea, tenha sido introduzido na Península Ibérica pelos Romanos, que denominavam *lactuca*, em Portugal utiliza-se ainda hoje o termo de origem árabe, enquanto em espanhol, pelo contrário, emprega-se o item lexical *lechuga* que se originou, com efeito, da respetiva palavra latina. Contudo, nas terras lusófonas circulava, antes da chegada dos povos muçulmanos, o termo *leituga* de origem românica que, provavelmente, foi substituído por *alface* na prática comunicativa, através do rápido processo de arabização. Diante desta particular situação linguística, é fundamental, por conseguinte, tentar perceber as razões que determinaram a sobrevivência secular do termo de origem árabe *alface*, registado no século XVI, em detrimento da

10 Machado 1956–59: s. v. *melão*. O lema *bateca* foi registado no século XVI (cfr. Idem: s. v. *bateca*), informação cronológica que explica o respetivo desaparecimento na dimensão lexical, uma vez que o uso comunicativo do sinónimo de origem latina *melão* era já enraizado no espaço lusófono a partir do século XI.

11 O verbe *figo* foi registado no século XIII, enquanto não há dados relativamente ao período de atestação do arabismo *xarifa* (veja-se Machado 1956–59: s. v. *figo* e *xarifa*).

respetiva unidade lexical de proveniência latina, já existente antes do século VIII e atestada, como informa Machado, no século XV (1956–59: s. v. *leituga*).

Para encontrar uma resposta à questão colocada, vale analisar a definição semântica de cada lema: *Infopédia* refere que a *alface* é uma “planta herbácea comestível, da família das Compostas, muito cultivada em Portugal, cujas folhas, bem como as de outras variedades da mesma espécie, são muito utilizadas em culinária, sobretudo na preparação de saladas”, enquanto a *leituga* é a “designação comum, extensiva a diferentes plantas herbáceas de seiva leitosa, da família das Compostas, que inclui espécies espontâneas em Portugal”. Levando em conta estas indicações, podemos aventar a hipótese de que o termo *leituga*, mais antigo, foi utilizado em função de hiperónimo para indicar a família botânica de pertença da hortaliça, enquanto o arabismo *alface*, empregado pelos falantes como hipónimo, indica a tipologia de planta pertencente à família das Compostas que se cultiva especialmente em terra portuguesa. Diante disto, podemos afirmar que durante a fase de bilinguismo românico-árabe (Suisse 2020: 47), existia no uso comunicativo o par sinonímico *alface/leituga*: o termo de origem latina era utilizado, segundo o que parece, para expressar um significado genérico, enquanto o arabismo, mais pragmático no domínio interativo, servia para indicar os traços sémicos específicos do referente (Santos 1980: 587), etiqueta lexical de maior emprego na altura porque, como já indicámos anteriormente, utilizar palavras de origem árabe na época de dominação islâmica era índice de prestígio cultural, atitude que fez circunscrever, por conseguinte, o emprego da palavra românica anterior, com o mesmo significado, ao âmbito da botânica. Além disso, este verbete consolidou-se de maneira tão intensa nos hábitos comunicativos da área portuguesa, para identificar esta tipologia de hortaliça que crescia de maneira luxuriante sobretudo na área lisboeta, que surgiu a alcunha *alfacinhas* para identificar os seus habitantes, definidos como “amigos de alfaces por as comerem exageradamente” (Reis 1943: 210). Como informa Machado no seu dicionário (1956–59: s. v. *alface*), este termo foi registado pela primeira vez no século XIX pelo escritor romântico João Baptista da Silva Leitão de Almeida Garrett (1799–1854) em *Viagens na minha terra*¹², de molde a diferenciar os lisboetas dos *saloios*, os camponeses da periferia (Carvalho 2006: 90), epíteto igualmente de origem árabe existente apenas no léxico português e ainda em uso em época hodierna.

Outro caso extremamente significativo de par sinonímico, entre as unidades lexicais antigas atestadas no século XII, que sobreviveram às mudanças diacrónicas da língua portuguesa, concerne o uso ainda atual dos arabismos *teiga* (existente também em espanhol) e *alcofa* (peculiar do espaço lusófono) – registados, segundo Machado (1956–59: s. v. *teiga* e *alcofa*), respetivamente em 1175 e 1188 – dois significantes que aludem ao mesmo referente, mas com características materiais diferentes: de facto, ambos identificam um cesto, realizado com palha no primeiro caso e com vime, palma ou esparto no segundo. Estes termos de origem arábica entraram no uso comunicativo das comunidades ibéricas apesar da anterior existência, como defende Maria José de Moura Santos (1980: 587), da forma românica *cesto* (proveniente do latim *caestu*), circunstância que originou quer uma situação de bilinguismo românico-árabe (Suisse 2020: 47), quer o uso paralelo de significantes de diferente etimologia, mas com traços sémicos comuns: podemos portanto supor que, nesta fase de evolução lexical, tal como aconteceu com o par *alface/leituga*, o termo neolatino era utilizado com função de hiperónimo para expressar, como sugere Santos (1980: 587), um

12 Garrett, J. B. D. A. (1856). *Viagens na minha terra*. Lisboa: Typ. Gazeta dos Tribunais, p. 61.

significado mais genérico, enquanto o arabismo, em termos de hiponímia, servia para indicar elementos sémicos específicos, atitude linguística que perdurou até aos nossos dias. Neste caso particular, *teiga* servia para designar – tal como hoje – um cesto de palha de pequenas dimensões, enquanto *alcofa*, pelo contrário, identificava um cesto de tamanho maior em vime com asa, etiquetas lexicais necessárias na altura para diferenciar todas as tipologias de cesto existentes em terra lusitana depois do processo de islamização.

A presença de arabismos peculiares foi detetada também no campo conceitual do vestuário, já que termos como *fato*¹³ – o item mais utilizado segundo os dados conseguidos no levantamento – *algibeira* (de média frequência) e *ceroulas* (de módico emprego) constituem a base lexical do português contemporâneo. O termo *fato*, que designa um vestuário quer masculino, constituído por calças, casaco e colete, quer feminino, composto geralmente de casaco e saia, ou, em alternativa, calças ou vestido (Infopédia: s. v. *fato*) é mais utilizado em relação à respetiva unidade lexical de origem latina *traje*, existente também em espanhol: de facto, Corominas (1987: s. v. *traje*) refere que o lema castelhano, atestado no século XV, deriva da palavra homógrafa portuguesa – registada, todavia, no mesmo período (Machado 1956–59: s. v. *traje*) em área lusitana – circunstância que faz intuir que o item entrou no uso comunicativo em Espanha através do vizinho idioma. Enquanto no território castelhano emprega-se *traje*, no espaço lusófono, pelo contrário, resulta mais frequente o uso do equivalente semântico *fato*, com 239.097 ocorrências, enquanto o sinónimo *traje* tem uma frequência de uso mais reduzida, recorrendo 4.776 vezes nos textos do corpus eletrónico examinado. Consultando a definição dada pelo dicionário online Priberam dos dois termos (s. v. *fato* e *traje*), resulta que o arabismo indica o “conjunto das peças de roupa”, enquanto o item de origem latina designa o “vestuário habitual”: com base nisto, podemos afirmar que, também para este caso de par sinónimo, o lema românico é utilizado como hiperónimo para identificar, em termos gerais, a indumentária habitual, enquanto *fato*, desempenhando uma função de subordinação conceptual, é o hipónimo que possui elementos sémicos específicos, referindo-se a uma combinação de peças distintas de vestuário para ambos os sexos.

Os arabismos *algibeira* – empregado nomeadamente nas regiões do Alentejo e do Algarve (Santos 1980: 590) – e *ceroulas*, atestados ambos no século XVI, resultam ser utilizados, contrariamente, cada vez menos, com 171 e 80 ocorrências contabilizadas: na prática discursiva habitual, emprega-se de maneira maioritária o sinónimo de origem latina *bolso*, lema que recorre nos textos do corpus Now 23.099 vezes, enquanto o verbete *ceroulas*, espécie de collantes usados antigamente pelos homens, caiu progressivamente em desuso, uma vez que esta peça de vestuário deixou de se usar.

Adotadas à chegada dos dominadores muçulmanos pelos falantes ibéricos, estas etiquetas estão a desaparecer progressivamente no português hodierno por uma questão de economia linguística, tal como aconteceu com os arcaísmos do mesmo setor conceitual – respetivamente *abá* (manto largo utilizado em lugares orientais), *alfola* (túnica de seda), *algebeta* (tipologia de gabão), *aljaravia* (túnica comprida com capuz), *bedém* (capa oriental), *cerame* (capa com capuz), *enxaravia* (faixa de cabelo ornamental feminina), *falifa* (roupa de couro), *fota* (turbante mourisco) e *tarbuche* (barrete turco para homens) – que resultaram ativos apenas durante o período de bilinguismo luso-árabe para identificar os novos hábitos indumentários.

13 Apesar de existir em espanhol o termo *hato*, igualmente de origem árabe, incluímos o lema *fato* nos arabismos peculiares da língua portuguesa, uma vez que, no espaço lusófono, o verbete indica um conjunto de peças de vestuário (Priberam: s. v. *fato*), enquanto nos falares hispânicos corresponde à indumentária cotidiana (RAE: s. v. *hato*).

O fenómeno da importação lexical foi fecundo também no âmbito semântico profissional, uma vez que muitos termos da língua de adstrato enriqueceram e renovaram o domínio verbal do setor político, administrativo, económico e militar, campos em que foram introduzidas, com efeito, noções e cargos laborais com aceções denotativas inéditas. Se consultarmos os dados conseguidos através do levantamento, é possível afirmar que, nesta categoria conceitual, domina o emprego comunicativo de 4 arabismos peculiares, i. e. 2 de alta frequência – *algez* (sinónimo de verdugo), cujo uso é ainda elevado por ter sido atestado no século XIV, antes do respetivo lema de origem latina¹⁴, e *alicerce* (fundamento de uma construção) –, 1 de média frequência – *algema* (o objeto metálico composto por duas argolas utilizado pelos agentes de polícia para prender alguém pelos pulsos) – e 1 de baixo uso – *enxovia* (espécie de prisão subterrânea) – cuja aceção denotativa originária ainda se preserva nos atuais hábitos comunicativos. Vale salientar, contudo, que muitas etiquetas lexicais deste campo semântico caíram em desuso, sendo ausentes nos textos do corpus *Now* – tal como *ababe* (marinheiro), *abaci* (moeda circulante na antiga Pérsia e em Índia), *abraemo* (moeda da antiga colónia goesa), *açaria* (incursão), *açalmo* (pagamento antecipado de provisões ou escravos¹⁵), *açougagem* (uma tipologia de taxa alimentar), *adil* (revendedor), *adiceiro* (hoje substituído pelo lema de origem latina *mineiro*, apesar de ter sido atestado posteriormente¹⁶), *alamia* (armadura), *alfândega* (sinónimo de *aduana*, outro arabismo empregado também em espanhol¹⁷), *alfitra* (imposto que os Muçulmanos tinham que pagar aos Cristãos), *alforma* (salvo-conduto), *algibebe* (revendedor), *anadel* (capitão de besteiros), *anoque* (lugar onde se trabalhava o couro), *anovel* (desaparecido em favor de *besteiro*¹⁸), *aqueme* (governador muçulmano), *baar* (tipologia de medida indiana, hoje mais utilizado como topónimo), *bazaruco* (moeda circulante na Índia portuguesa), *caçabé* (castelo¹⁹ XII), *cafarro* (imposto pago em terra santa), *catual* (funcionário público oriental), *deviva* (*alfândega*), *faraçala* (uma medida de peso), *faraz* (estofador), *haquim* (mestre, professor), *lascari* (marinheiro indiano), *marzagani* (cavaleiro dos territórios de língua árabe), *mocarraria* (espécie de taxa), *molana* (dono, proprietário), *názir* (superintendente das mesquitas), *para* (moeda turca), *ságena* (tipologia de prisão) – que desapareceram, segundo o nosso parecer, quer por uma questão de economia e de redundância linguística, quer por identificarem conceitos cultural e geograficamente circunscritos, inexistentes nos territórios lusófonos, que se empregavam nos falares ibéricos apenas no período de ampla difusão, através da intermediação linguística dos Moçárabes, do idioma árabe.

14 O lema *verdugo* foi atestado, como refere Machado, no século XVI, enquanto o arabismo *algez* duzentos anos antes (cfr. Machado 1956–59: s. v. *algez* e *verdugo*).

15 Alves 2013: s. v. *açalmo*.

16 Com base nas informações etimológicas consultadas, o verbete *adiceiro* caiu em desuso embora tenha sido atestado em 1325, ao passo que o respetivo sinónimo *mineiro*, registado no século XVII, é atualmente uma entrada de alta frequência (cfr. Machado 1956–59: s. v. *adiceiro* e *mineiro*).

17 Apesar do termo *alfândega* ter sido atestado em 1249, i.e. anteriormente em relação à *aduana*, registado ao invés no século XVI, tornou-se num arcaísmo. Segundo os dados identificados através do levantamento realizado, o lema parece ter caído em desuso, embora seja palavra corrente em PE: supomos que o número de ocorrências é baixo por ser um termo empregado nomeadamente em contextos socio-discursivos específicos.

18 *Besteiro* era um termo já consolidado no uso comunicativo dos falares ibéricos, uma vez que foi registado em 1360, enquanto o arabismo *anovel* remonta ao século XV (cfr. Machado 1956–59: s. v. *anovel* e *besteiro*).

19 O lema *caçabé*, atestado no século XVI, tornou-se num arcaísmo, dado que o sinónimo de origem latina *castelo* já era utilizado na prática verbal em 1142 (veja-se Machado 1956–59: s. v. *caçabé* e *castelo*).



Além disso, podemos afirmar que também o campo semântico respeitante às mercadorias e aos utensílios, introduzidos nos hábitos cotidianos, resulta ser muito prolífico em termos de importação lexical da *algaravia*. De facto, no português resultam ativos, ainda hoje, 5 termos pertencentes a este setor conceitual, como *alguidar* (uma espécie de vaso de barro ou madeira), *cabide* (peça de madeira ou plástico para pendurar peças de vestuário), *gazua* (chave falsa), *giz* (substância sólida utilizada para escrever sobre quadros) e *sucata* (ferro-velho). Obviamente, durante o período de islamização, entraram numerosos itens em relação a esta categoria semântica, mas que desapareceram gradualmente da linguagem corrente por serem substituídos nomeadamente por etiquetas de origem latina, como *açafeia* (ardósia²⁰), *açação* (tipo de material quente), *açamo* (variante de *açaime*, mais utilizada hoje na língua corrente), *adival* (corda de carro), *agomia* (tipo de cimitarra), *albafar* e *albafor* (termos desaparecidos em favor de *incenso*²¹), *alcaiatão* (tenda empregada em terras orientais), *alcaima* (tenda de lã utilizadas pelos povos nómades), *alcânave* (variedade de linho), *alcaraviz* (tubo de ferro que leva o ar do fole para a forja), *alcatrate* (prancha utilizada nas naves), *alcatruz* (vaso de barro), *alchaz* (espécie de pano), *alfaia* (utensílio de lavoura), *alfarrábio* (deantropónimo atestado no século XVIII que se originou do nome do filósofo Al-Farâbi²²), *alferça* (tipo de picareta), *alfrezes* (espécie de cobertor), *algeroz* (tipo de caleira), *aljorce* (arabismo que caiu em desuso em favor de *chocalho*²³), *aljuz* (cola extraída da planta do cardo), *almáfega* (sinónimo de cabeceira²⁴), *almafre* (tipologia de capacete utilizado pelos soldados), *almandra* (coberta de cama), *almanjarra* (pau para segurar o animal à nora), *almanxar* (terraço onde se secavam os figos), *almarcova* (espécie de cutelo), *almatrixa* (almofada para cobrir o albardão), *almofrez* (termo desaparecido e substituído por *furador*, já em uso a partir do século XIII²⁵), *almotolia* (vaso para azeite), *alparavaz* (espécie de franja), *alpe* (género de almofada), *alquitão* (tipo de tenda), *altâncaro* (pandeiro mourisco), *anafia* (trombeta de metal), *atabaque* (categoria de tambor), *atafera* (cinta de esparto), *baraço* (cordão²⁶), *belmaz* (variante ortográfica de *alfinete*), *choca* (tipo de bola), *chúmeas* (peças para ligar os mastros), *cifa* (areia para moldar), *delu* (tipo de vaso), *enxarrafa* (cordão de seda), *fateixa* (espécie de âncora com várias hastes para pequenas embarca-

20 Tendo em consideração as informações etimológicas consultadas, ainda não se conhece o ano de registo do arabismo, enquanto o termo de origem latina *ardósia* foi atestado em 1871 (veja-se Machado 1956–59: s. v. *açafeia* e *ardósia*).

21 A entrada de origem latina *incenso* já era utilizada pelos povos neolatinos a partir do século XIII, enquanto os arabismos *albafar* e *albafor* foram registados, respetivamente, nos séculos XIX e XVII (veja-se Machado 1956–59: s. v. *albafar*, *albafor* e *incenso*).

22 Alves 2013: s. v. *alfarrábio*.

23 O termo *chocalho*, de origem latina, já se empregava no espaço lusófono desde o século XIII, enquanto o arabismo *aljorce* foi atestado posteriormente, no século XIX, circunstância que explica a extinção desta palavra (cfr. Machado 1956–59: s. v. *aljorce* e *chocalho*).

24 Apesar do arabismo ter sido atestado em 1369, hoje em dia utiliza-se o sinónimo de origem latina *cabeceira*, registado no século XV (veja-se Machado 1956–59: s. v. *almáfega* e *cabeceira*).

25 Enquanto não se conhece ainda o período de atestação do arabismo *almofrez*, pressupomos que caiu em desuso porque o sinónimo de origem latina *furador* já se utilizava nos falares ibéricos durante o século XIII (cfr. Machado 1956–59: s. v. *almofrez* e *furador*).

26 *Baraço* e *cordão* entraram ambos no uso comunicativo no século XIII, mas o arabismo, segundo os dados do levantamento (19 ocorrências), tornou-se num arcaísmo (veja-se Machado 1956–59: s. v. *baraço* e *cordão*). Apesar do arabismo ter caído em desuso, alguns derivados fazem parte da língua corrente, tal como *embarçar* e *desembarçado* que apresentam respetivamente 462 e 98 ocorrências.

ções, termo utilizado sobretudo no sul de Portugal), *fazançal* (tecido produzido na Pérsia), *gelba* (pequena embarcação), *manchil* (género de faca) e *rusma* (creme depilatório usado no Oriente).

O fenómeno do empréstimo lexical do árabe verificou-se também no âmbito rural, sendo muitas as inovações técnicas, tal como a famosa *nora*, que os islâmicos introduziram na Península Ibérica. Os arabismos desta categoria, que ainda se utilizam no léxico contemporâneo, são apenas 2, i. e. *alcateia* (grupo de animais ferozes) e *ceifa* (colheita de cereais), ambos de média frequência, enquanto as outras etiquetas verbais tornaram-se em arcaísmos, termos avulsos que se empregaram unicamente durante a fase de bilinguismo luso-árabe, como: *adil* e *alqueive* (ambos extintos em favor de *pousio*, já em uso a partir do século XIV²⁷), *aguedal* (tipo de terreno cultivado), *alabão* (sinónimo de *gado*²⁸), *alacil* (período da estação da vindima para os Mouros), *alcanchal* (caminho intransitável), *alchazar* (campo de forragem), *alcórcova* (fosso), *alfeire* (curral de porcos), *alferga* (medida para o milho), *alfobre* (viveiro de plantas), *alfoufe* (lote de terreno), *algame* (trabalhador rural), *almoqueire* (condutor de bestas), *asusada* (tomada de água), *azarve* (proteção contra o vento), *azinhaga* (caminho estreito), *enxaca* (bolsa para animais de carga), *enxerca* (a operação de cortar a carne para secá-la ao sol²⁹), *lezira* (variante de *lezíria*, terreno nas margens dum rio), *manjorra* (travessa para segurar os animais à nora) e *maracha* (terreno alagadiço).

Quanto aos hábitos socioculturais, pelo contrário, os lemas de origem islâmica ainda ativos são 6, i.e. *aldrabão* (pessoa que mente), *anafado* (gordinho), *algarvio* (oriundo da região do Algarve, com 4.663 ocorrências), *arraia* (populacho), *arraia-miúda* (populacho) e *azáfama* (sinónimo de pressa). A maior parte destes empréstimos caiu em desuso, uma vez que tais unidades lexicais indicavam paradigmas culturais típicos das tradições muçulmanas, que desapareceram no território ibérico, de facto, com o fim da Reconquista. Citamos, a tal propósito, os arcaísmos *ababil* (ave citada no Alcorão), *abaná* (nome de uma tribo), *açacalar* (limpar armas brancas), *acedrenche* (variante ortográfica de *xadrez*), *açougui* e *alcaçaria* (praça onde se realizava o mercado), *alaela* (multidão de pessoas), *alcaiotaria* (ato de mexericar), *alcaiote* e *alcofa* (rufião³⁰), *alcarrada* (jogo de origem oriental), *alcoceifa* e *alcouce* (prostíbulo para os povos orientais), *alguergue* (antigo jogo árabe), *alhelda* (tipo de acampamento), *almojama* (ponto de encontro), *azenegue* (língua bérbere), *carava* (malta), *fona* (pessoa avarenta), *iben* (filho), *marafona* (meretriz) e *sáfaro* (pessoa esquiva).

Relativamente à área científica, os resultados do levantamento mostraram que apenas o verbete *adelbarão*, de baixa frequência, pertence ao léxico português hodierno, enquanto todas as palavras que, a partir do século XVIII, entraram para designar todas as inovações introduzidas pelos Muçulmanos neste setor, desapareceram progressivamente no uso comunicativo, tal como *alcouço* (território do Sul), *alcovês*, *camsim* e *xamal* (tipologias de vento), *alfarde*, *alfisque*, *algenibe*, *algorabe*, *alquês*, *altair*, *aludra* e *cotamaluco* (nomes de estrelas), *alfridária* (a influência dos astros),

27 Quanto a *adil*, é desconhecido o período de atestação; *alqueive*, pelo contrário, remonta ao século XVI (cfr. Machado 1956–59: s. v. *adil* e *alqueive*).

28 *Gado* utilizava-se nos territórios lusófonos da Península Ibérica já a partir de 837, enquanto *alabão*, verbe hoje extinto, foi atestado no século XVII (Machado 1956–59: s. v. *alabão* e *gado*), provavelmente durante o terceiro período de entrada dos arabismos (Suisse 2020: 47) através do contato diuturno entre a língua portuguesa e os falares orientais nas colónias extracontinentais.

29 Alves 2013: s. v. *enxerca*.

30 Segundo os dados até hoje disponíveis, ainda resulta desconhecido o período de atestação dos dois arabismos *alcaiote* e *alcofa*, enquanto o correspondente termo de origem latina *rufião* foi atestado no século XV (cfr. Machado 1956–59: s. v.: *alcaiote*, *alcofa* e *rufião*).

almucábala (variante de *álgebra*), *almucantarado* (relativo ao *Almucantarat*), *azinhavre* (variante de *zinabre*), *nochatro* (sal amoníaco) e *tutanagá* (liga de estanho).

No âmbito da religião, os 2 arabismos ainda em uso são respetivamente *haje* (peregrino muçulmano) e *mafarrico* (anjo decaído que procura afastar o fiel de Alá³¹), enquanto os restantes 13 lemas já não se utilizam na prática verbal contemporânea porque se referem à dimensão específica do islamismo: i. e. *açómua* (minarete³²), *açubá* (oração das comunidades pretas do Brasil), *ais-saua* (tipo de confraria), *alimangariba* (tipo de oração), *almeazar* (toalha para o altar), *almocávar* (cemitério mouro), *alufá* (religioso muçulmano do oriundo do norte de África), *caciz* (sacerdote árabe), *calândar* (monge maometano), *celá* (tipo de oração), *moçafo* (volume do Alcorão), *rafidi* (xiita) e *sai* (bonzo).

O fenómeno de importação lexical, obviamente, manifestou-se também na área da flora e da fauna, à qual pertencem os arabismos *alforreca* (sinónimo da palavra de origem latina *medusa*, mais empregada do ponto de vista comunicativo³³) e *osga* (pequeno sáurio)³⁴, ambos de baixa frequência, uma vez que recorrem respetivamente 69 e 21 vezes nos textos digitais do corpus eletrónico *Now*. Entre as unidades verbais que desapareceram dos hábitos linguísticos mencionamos: *abelmeluco* (rícinio³⁵), *albafar* (tipologia de tubarão), *alcabroz* (peixe vulgar no Algarve), *alcaravão* (tipo de ave), *alfario* (burro selvagem), *bejabã* (deserto), *benefe* (violeta-brava), *dorónico* (tipo de flor), *harmala* (arruda silvestre), *meari* (tipologia de dromedário) e *nilgó* (espécie de antílope).

Para terminar este panorama lexical, baseado numa abordagem tanto diacrónica como sincrónica, vale a pena elencar, a título elucidativo, os arabismos que pertencem ao setor das expressões e de várias áreas semânticas que, mesmo sendo parte fundamental do património histórico linguístico do português, extinguíram-se com o gradual desaparecimento da cultura islâmica do território lusitano ou, por uma questão de economia linguística, foram substituídos por lemas de origem latina, já em uso antes do período secular de dominação muçulmana. Enquanto *taibo* (agradável) resulta ser parcamente utilizado, as outras expressões e palavras identificadas extinguíram-se: *cor aberta* (cor clara) e *cor fechada* (cor escura), *afagamento* e *tagaté* (carícia³⁶), *ala!* (Força!), *alcarrada* (movimento das aves de rapina), *alfétena* (guerra civil), *alfinago* (pessoa canalha), *alforra* (erupção cutânea), *alfurja* (espécie de pátio interior), *almajuz* (alcunha para identificar os Normandos), *ateigar* (avaliar a olho), *aziz* (ilustre), *beliz* (sagaz), *bissimilai* (“em nome de Alá”), *cadimo* (cujo significado é ladrão), *cecotrim* (oriundo da ilha de Socotorá), *fartaquim*

31 Alves 2013: s. v. *mafarrico*.

32 Segundo as informações etimológicas disponíveis, ainda resulta desconhecido o período de registo dos dois verbetes (cfr. Machado 1956–59: s. v. *açómua* e *minarete*).

33 O termo *medusa*, como refere Machado, foi atestado no século XVI, enquanto o arabismo entrou no uso comunicativo no ano de 1712 (cfr. Machado 1956–59: s. v. *alforreca* e *medusa*). Esta circunstância explica o desaparecimento gradual do arabismo na prática conversacional, uma vez que o verbete de origem latina foi registado anteriormente.

34 O lema *osga*, sendo um termo pertencente ao campo semântico da fauna, não recorre frequentemente nos textos do corpus, uma vez que se utiliza nomeadamente em contextos socio-discursivos específicos.

35 As duas entradas foram atestadas ambas no século XIX mas, não obstante, o arabismo caiu em desuso (cfr. Machado 1956–59: s. v. *abelmeluco* e *rícinio*).

36 O lema *afagamento* foi registado no século XIII, enquanto o sinónimo *tagaté* no século XIX. Ambos caíram em desuso, uma vez que, provavelmente, o verbete de origem latina *carícia* já se utilizava na prática comunicativa antes da chegada dos Muçulmanos, apesar de ter sido atestado pela primeira vez no século XVII (cfr. Machado 1956–59: s. v. *afagamento*, *tagaté* e *carícia*).

(chefe de Fartak), *lobão* (sinónimo de *benjoim*, também este um arabismo), *moxinifada* (mistura), *mozmodi* (tribo bérbere), *muaxa* (tipologia de verso literário utilizado no Oriente), *muálaca* (hesitante), *nacibo* (nobre), *nafar* (pessoa), *odiá* (presente), *olede* (tipo de tribo), *oução* (ácara que se encontra no queijo), *valide* (recém-nascido), *xué* (coisa de pouco valor), *zirbo* (peritoneu) e *zuaque* (termo popular para indicar o traseiro).

A título conclusivo, visando uma melhor análise do tema proposto, é importante indicar como o significado denotativo de alguns arabismos analisados mudou complementemente em perspetiva diacrónica, já que 23 arcaísmos adquiriram, tendo em consideração a análise qualitativa das ocorrências, a função de i) topónimos pertencentes a áreas orientais, portuguesas e brasileiras – como no caso de *Alcouce*, *Alferce*, *Azinhaga*, *Baar* e *Fateixa* –, ii) antropónimos – *Adel*, *Adil*, *Altair*, *Azinhaga* e *Aziz* – e iii) apelidos, i. e. *Alamia*, *Alfaria*, *Alforra*, *Aziz*, *Beliz*, *Buza*, *Faraz*, *Fateixa*, *Fona*, *Fota*, *Lobão*, *Marafona* e *Xué*.

4. Considerações finais

Ao concluirmos a análise apresentada – baseada na contabilização das ocorrências de cada arabismo próprio da língua portuguesa, atestadas nos textos do corpus eletrónico *Now*, e na comparação entre a data de atestação dos lemas constituintes os pares sinonímicos latino-árabe – advogamos a tese de que a maior parte das unidades lexicais que entraram no territórios do atual Portugal, a partir do século VIII, caiu em desuso, uma vez que apenas 39 lemas sobre 267 resultam ainda ativos no português contemporâneo, dos quais só 10 pertencem à categoria de alta frequência – *fato* (239.097), *fatia* (15.069), *algarvio* (4.663), *alface* (3.089), *sucata* (2.559), *algoz* (2.368), *alicerce* (1.365), *giz* (1.348), *cabide* (831) e *azáfama* (801) –, 13 são de média frequência – *arraia* (469); *alcatra* (454); *açorda* (439); *algema* (415); *alcateia* (238); *saloio* (228); *aldabrão* (212); *regueifa* (210); *algibeira* (171); *alguidar* (170); *ceifa* (161); *acém* (144) e *alqueire* (108) –, e 16 de baixa frequência, respetivamente *ceroulas* (80), *alforreca* (69), *adiafa* (68), *gazua* (55), *haje* (52), *alcofa* (47), *aluá* (44), *frangão* (30), *enxovia* (37), *anafado* (38), *mafarrico* (32), *taibo* (27), *arraia-miúda* (23), *aldebarã* (21), *osga* (21) e *azurracha* (termo empregado unicamente nas regiões setentrionais de Portugal).

Através do exame conduzido, verificámos que grande parte do património lexical de origem árabe – 85 % – era utilizada no âmbito comunicativo principalmente durante o período de ocupação islâmica do espaço lusófono da Península Ibérica, onde, com a entrada da língua de adstrato, foram assimiladas numerosas etiquetas verbais para denotar quer i) as inovações no âmbito científico, rural e profissional introduzidas pelos dominadores, quer ii) aspetos da vida sociocultural historicamente circunscritos, como os vários termos adotados na área semântica da religião muçulmana, do vestuário e do campo da administração política que desapareceram gradualmente com o fim da dominação islâmica. Com base nesta constatação, podemos afirmar, com efeito, que os pares sinonímicos latino-árabe ainda existentes permaneceram no desempenho comunicativo hodierno para distinguir referentes com entidades sémicas específicas ou significação geral, como, a título elucidativo, no caso de *leituga* e *alface*.

A realização deste estudo tanto qualitativo como quantitativo demonstrou, de facto, que os restantes 228 arabismos tornaram-se em arcaísmos, uma vez que alguns destes itens foram substituídos

pelos respetivos sinónimos de origem latina que, como apurámos através da colação das datas de registo fornecidas nomeadamente pelo dicionário etimológico de José Pedro Machado (1956–59), já se utilizavam anteriormente ao século VIII nos falares ibéricos: esta circunstância explica, por conseguinte, a gradual extinção de vários lemas, símbolo autêntico de um valioso passado histórico, sob o ponto de vista cultural, da dimensão lexical diacrónica da língua portuguesa.

Referências bibliográficas

- Alves, A. (1999). *Portugal: ecos de um passado árabe*. Lisboa: Instituto Camões.
- . (2013). *Dicionário de Arabismos da Língua Portuguesa*. Lisboa: Imprensa Nacional-Casa da Moeda.
- Biderman, M. T. C. (1978). *Teoria Linguística: Linguística Quantitativa e Computacional*. Rio de Janeiro: Livros Técnicos e Científicos.
- . (1998). A face quantitativa da linguagem: um dicionário de frequências do português. *Alfa*, 42, 161–181.
- Buescu, M. L. C. (1983). *O Estudo das Línguas Exóticas no Século XVI*. Lisboa: Instituto de Cultura e Língua Portuguesa.
- Carvalho, F. A. (2006). O lugar dos negros na imagem de Lisboa. *Sociologia, Problemas e práticas*, 52, 87–108.
- Castro, I. (1991). *Curso de história da língua portuguesa*. Lisboa: Universidade Aberta.
- Corominas, J. (1954). *Diccionario Crítico Etimológico de la Lengua Castellana*. Berna: Francke.
- Corriente, F. (2003). *Diccionario de arabismos y voces afines en iberorromance*. Madrid: Gredos.
- Cunha, A. G. D. (2011). *Dicionário etimológico da língua portuguesa*. Rio de Janeiro: Lexikon Editora Digital Ltda.
- Farinha, A. D. (1973). Contribuição para o estudo das palavras portuguesas derivadas do árabe hispânico. In *Portugaliae Historica VI* (pp. 244–265). Lisboa: Faculdade de Letras da Universidade de Lisboa.
- Franca, R. (1994). *Arabismos: uma minieniclopédia do mundo árabe*. Recife: EDUFPE.
- Guichard, P. (1977). *Structures sociales «orientales» et «occidentales» dans l'Espagne musulmane*. Paris: La-Haye.
- Lopes, D. (1968). *Nomes árabes de terras portuguesas*. Lisboa: Sociedade da Língua Portuguesa.
- Lüdke, M.; & André, M. (1986). Abordagens qualitativas de pesquisa: a pesquisa etnográfica e o estudo de caso. In M. Lüdke, & M. André (Eds.). *Pesquisa em educação: Abordagens qualitativas* (pp. 11–22). São Paulo: EPU.
- Machado, J. P. (1940). *Sintra muçulmana*. Sintra: Imprensa Nacional.
- . (1956–1959). *Dicionário etimológico da língua portuguesa*. Lisboa: Editorial Confluência. 5 voll.
- . (1958). *Influência Árabe no Vocabulário Português*. Lisboa: A. Pinto.
- . (1991). *Vocabulário Português de Origem Árabe*. Lisboa: Editorial Notícias.
- Machado Filho, L. (2013). Arabismos e germanismos em textos medievais portugueses. In *Encontro Internacional dos Estudos Medievais – ABREM 10* (pp. 392–400). Brasília: Editora Universidade de Brasília.
- Maranhão, S. M. (2018). Arabismos portugueses no contexto multilinguístico da Península Ibérica Medieval. *Calígrama*, 23, 2, 121–143.
- Mattoso, J. (1985). Os Moçárabes. *Revista Lusitania*, 6, 5–24.
- Picard, C. (1983). Les mozarabes dans l'Occident ibérique (VIII-XII siècle). *R.E.I.*, 51, 77–88.

- . (2000). *Le Portugal musulman (VIII-XIII siècle). L'Occident d'al-Andalus sous domination islamique*. Paris: Maisonneuve et Larose.
- Pinto, M. D. C. (2009). O Douro no Gharb al-Ândalus: a história e a arquitectura do Douro entre os séculos VIII e XII. *A obra nasce: revista de Arquitectura da Universidade Fernando Pessoa*, 6, 53–64.
- Reis, F. (1943). Alfacinhas. *Olisipo, Boletim do grupo «Amigos de Lisboa»*, 24, 210–214.
- Ribeiro, O. (1965). A propósito de áreas lexicais no território português (Algumas reflexões acerca do seu condicionamento). *Separata do Boletim de Filologia*, XXI, 3–4, 177–208.
- Rossi, M. A. (2011). Dall'Africa al Portogallo: gli arabismi peculiari della lingua portoghese. In *Viagens de línguas e culturas* (pp. 129–163). Lisboa: 101 Noites.
- Sousa, J. D. (1830). *Vestígios da língua árabe em Portugal*. Lisboa: Academia Real das Sciencias.
- Samu, L. (2010). Presença árabe no português: 1300 anos depois. *Revista Augustus*, 30, 46–51.
- Santos, M. J. D. M. (1980). Importação lexical e estruturação semântica: os arabismos na língua portuguesa. *Biblos*, 56, 537–598.
- Suisse, A. (2020). Aspetos importantes do legado árabe na língua portuguesa. *Revista Língua-lugar*, 2, 43–59.
- Teysier, P. (1982). *História da Língua Portuguesa*. Lisboa: Sá da Costa.
- Turato, E. R. (2004). A questão da complementaridade e das diferenças entre métodos quantitativos e qualitativos de pesquisa: uma discussão epistemológica necessária. In S. Grubits, & J. A. V. Noriega (Eds.). *Método qualitativo: epistemologia, complementariedades e campos de aplicação* (pp. 17–51). São Paulo: Vetor Editora.
- Vargens, J. B. M. (2007). *Léxico português de origem árabe: subsídios para os estudos de filologia*. Rio de Janeiro: Almadena.
- Vasconcelos, C. M. de. (1956). Fontes do léxico português: os elementos árabes. In *Lições de filologia portuguesa* (pp. 299–310). Lisboa: Revista de Portugal.
- Werlich, E. (1975). *Typologie der Texte; Entwurf eines textlinguistischen Modells zur Grundlegung einer Textgrammatik*. Heidelberg: Quelle & Meyer.
- Zaidan, A. (2010). *Letras e História: mil palavras árabes na língua portuguesa*. São Paulo: Escrituras Editora.



Anexo:Lista dos arabismos próprios da língua portuguesa e respetivas ocorrências no corpus *Now*

Arabismo	Ocorrências Corpus <i>Now</i>	Uso contemporâneo
Abá	3	Desuso
Ababe	0	Desuso
Ababil	3	Desuso
Abaci	1	Desuso
Abaná	0	Desuso
Abelmeluco	0	Desuso
Aberto (cor aberta)	6	Desuso
Abraemo	0	Desuso
Açacalar	0	Desuso
Açafeia	0	Desuso
Açafão	1	Desuso
Açalmo	0	Desuso
Açamo	5	Desuso
Açaria	0	Desuso
Acedrenche	0	Desuso
Acém	144	Média frequência
Açómua	0	Desuso
Açorda	439	Média frequência
Açougagem	0	Desuso
Açougui	0	Desuso
Açubá	0	Desuso
Adáçama	0	Desuso
Adel	798 como antropónimo, mas 0 como <i>revendedor</i> .	Desuso
Adiafa	68	Baixa frequência
Adiceiro	0	Desuso
Adil	274 como antropónimo, mas 0 como <i>pousio</i> .	Desuso
Adival	42 como antropónimo, mas 0 como <i>corda</i> .	Desuso
Adora	0	Desuso
Afaagamento	0	Desuso
Agomia	0	Desuso
Aguedal	0	Desuso
Aissaua	0	Desuso
Ala!	1	Desuso

Arabismo	Ocorrências Corpus Now	Uso contemporâneo
Alabão	0	Desuso
Alacil	0	Desuso
Alaela	0	Desuso
Alamia	4	Desuso
Albafar (1)	2 (tipo de tubarão)	Desuso
Albafar (2)	0 (incenso)	Desuso
Albafor	0	Desuso
Alcabroz	0	Desuso
Alcaçaria	0	Desuso
Alcaiatão	0	Desuso
Alcaima	0	Desuso
Alcaiotaria	0	Desuso
Alcaïote	0	Desuso
Alcânave	0	Desuso
Alcanchal	0	Desuso
Alcaraviz	0	Desuso
Alcarrada (1)	0 (jogo)	Desuso
Alcarrada (2)	0 (movimento das aves de rapina)	Desuso
Alcateia	238	Média frequência
Alcatra	454	Média frequência
Alcatrate	0	Desuso
Alcatruz	3	Desuso
Alcavarão	0	Desuso
Alchaz	0	Desuso
Alchazar	0	Desuso
Alcoceifa	0	Desuso
Alcofa (1)	47 (cesto)	Baixa frequência
Alcofa (2)	0 (rufião)	Desuso
Alcorcova	0	Desuso
Alcouce	36 como topónimo	Desuso
Alcouço	0	Desuso
Alcovês	0	Desuso
Alçuz	0	Desuso
Aldabarão/Aldebarã	0, Aldabarão / 21, Aldebarã	Baixa frequência
Aldrabão	212	Média frequência

Arabismo	Ocorrências Corpus <i>Now</i>	Uso contemporâneo
Alface	3.089	Alta frequência
Alfaco	0	Desuso
Alfaia	228, como apelido	Desuso
Alfândega	0	Desuso
Alfarde	0	Desuso
Alfario	0	Desuso
Alfarrábio	17	Desuso
Alfava	1	Desuso
Alfeire	1	Desuso
Alféloa	0	Desuso
Alferça/Alferce	0, Alferça / 162, Alferce como topónimo	Desuso
Alferga	0	Desuso
Alfétena	2	Desuso
Alfinago	0	Desuso
Alfisque	0	Desuso
Alfitra	0	Desuso
Alfobre	0	Desuso
Alfola	0	Desuso
Alforma	0	Desuso
Alforra	6, como apelido	Desuso
Alforreca	69	Baixa frequência
Alfoufe	0	Desuso
Alfrezes	0	Desuso
Alfridária	0	Desuso
Alfurja	0	Desuso
Alganame	0	Desuso
Algarvio	4.663	Alta frequência
Algaz	0	Desuso
Algema	415	Média frequência
Algenibe	0	Desuso
Algeroz	11	Desuso
Algibebe	1	Desuso
Algibeira	171	Média frequência
Algibeta	0	Desuso
Algorabe	0	Desuso

Arabismo	Ocorrências Corpus Now	Uso contemporâneo
Algoz	2.368	Alta frequência
Alguergue	0	Desuso
Alguidar	170	Média frequência
Alhela	0	Desuso
Alicerce	1.365	Alta frequência
Alimangariba	0	Desuso
Aljaravia	0	Desuso
Aljorce	0	Desuso
Aljuz	0	Desuso
Almáfega	0	Desuso
Almafne	0	Desuso
Almajuz	0	Desuso
Almandra	0	Desuso
Almanjar/Almanjarra	0, Almanjar / 4, Almanjarra	Desuso
Almanxar	0	Desuso
Almarcova	0	Desuso
Almatriza	0	Desuso
Almeazar	0	Desuso
Almocávar	0	Desuso
Almofrez	0	Desuso
Almojama	0	Desuso
Almoqueire	0	Desuso
Almotolia	6	Desuso
Almucábala	0	Desuso
Almucantarado	0	Desuso
Alparavaz	0	Desuso
Alpe	82, com sentido de montanha <i>monte</i>	Desuso
Alqueire	108	Média frequência
Alqueive	2	Desuso
Alquês	0	Desuso
Alquitão	0	Desuso
Altair	1.729, como antropónimo	Desuso
Altáncaro	0	Desuso
Aluá	44	Baixa frequência
Aludra	0	Desuso

Arabismo	Ocorrências Corpus <i>Now</i>	Uso contemporâneo
Alufá	6	Desuso
Anadel	0	Desuso
Anafa	2	Desuso
Anafado	38	Baixa frequência
Anáfega	0	Desuso
Anafia	0	Desuso
Anoque	5	Desuso
Anovel	0	Desuso
Aqueme	0	Desuso
Arraia	469	Média frequência
Arraia-miúda	23	Baixa frequência
Asusada	0	Desuso
Atafera	0	Desuso
Ateigar	0	Desuso
Atequipera	0	Desuso
Azáfama	801	Alta frequência
Azarve	0	Desuso
Azenegue	0	Desuso
Azevizinhos	0	Desuso
Azinhaga	301, como antropónimo e topónimo	Desuso
Azinhavre	1	Desuso
Aziz	2.713, como antropónimo e apelido	Desuso
Azurracha	0, utilizado apenas na zona do Douro	Uso regional
Baar	34, como topónimo	Desuso
Bácoro	12	Desuso
Baraço	19	Desuso
Batafaluga	0	Desuso
Bateca	0	Desuso
Bazaruco	0	Desuso
Bedém	0	Desuso
Bejabã	0	Desuso
Beliz	87, como apelido	Desuso
Belmaz	0	Desuso
Benefe	0	Desuso
Bissimilai	0	Desuso

Arabismo	Ocorrências Corpus Now	Uso contemporâneo
Buza	20, como apelido	Desuso
Cabide	831	Alta frequência
Caçabé	0	Desuso
Caciz	0	Desuso
Cadimo	0	Desuso
Cafarro	0	Desuso
Calândar	0	Desuso
Camsim	0	Desuso
Carava	15	Desuso
Catual	5	Desuso
Cecotrim	0	Desuso
Ceifa	161	Média frequência
Celá	0	Desuso
Cerame	1	Desuso
Ceroulas	80	Média frequência
Choca	0	Desuso
Chúmeas	0	Desuso
Cifa	0	Desuso
Cotamaluco	0	Desuso
Delu	0	Desuso
Deviva	0	Desuso
Dorónico	0	Desuso
Enxaca	0	Desuso
Enxaravia	0	Desuso
Enxarrafa	0	Desuso
Enxerca	0	Desuso
Enxovia	37	Baixa frequência
Falifa	0	Desuso
Faraçala	0	Desuso
Faraz	6, como apelido	Desuso
Farroupo	0	Desuso
Fartaquim	0	Desuso
Fateixa	23, como topónimo e apelido	Desuso
Fatia	15.069	Alta frequência
Fato	239.097	Alta frequência

Arabismo	Ocorrências Corpus <i>Now</i>	Uso contemporâneo
Fazanzal	0	Desuso
Fechado	1, “cor fechada”	Desuso
Fístico	0	Desuso
Fona	124, como apelido	Desuso
Fota	0	Desuso
Frangão	30	Baixa frequência
Gazua	55	Baixa frequência
Gelba	0	Desuso
Giz	1.348	Alta frequência
Haje	52	Baixa frequência
Haquim	0	Desuso
Harmala	0	Desuso
Iben	7	Desuso
Lascari	1	Desuso
Lezira	0	Desuso
Lobão	4.156, como apelido	Desuso
Maçal	13	Desuso
Mafarrico	32	Baixa frequência
Manchil	0	Cutelo
Manjorra	0	Desuso
Maracha	2	Desuso
Marafona	958, como apelido	Desuso
Marzagani	0	Desuso
Meari	4	Desuso
Moçafo	0	Desuso
Mocarraria	0	Desuso
Molana	0	Desuso
Moxinifada	1	Desuso
Mozmodi	0	Desuso
Muaxa	0	Desuso
Muálaca	0	Desuso
Nacibo	0	Desuso
Nafar	2	Desuso
Názir	0	Desuso
Nilgó	0	Desuso

Arabismo	Ocorrências Corpus Now	Uso contemporâneo
Nochatro	0	Desuso
Odiá	0	Desuso
Olede	0	Desuso
Osga	21	Baixa frequência
Ouçã	0	Desuso
Para	0	Desuso
Rafidi	0	Desuso
Regueifa	210	Média frequência
Rusma	0	Desuso
Sáfaro	0	Desuso
Ságena	0	Desuso
Sai	0	Desuso
Saloio	228	Média frequência
Sucata	2.559	Alta frequência
Tagaté	0	Desuso
Taibo	27	Baixa frequência
Tarbuche	0	Desuso
Tutanagá	0	Desuso
Valide	0	Desuso
Xamal	0	Desuso
Xarifa	1	Desuso
Xué	2, como apelido	Desuso
Zirbo	0	Desuso
Zuaque	0	Desuso



This work can be used in accordance with the Creative Commons BY-SA 4.0 International license terms and conditions (<https://creativecommons.org/licenses/by-sa/4.0/legalcode>). This does not apply to works or elements (such as images or photographs) that are used in the work under a contractual license or exception or limitation to relevant rights.

