

# Estruturas árabes no léxico ibero-românico<sup>1</sup>

## Arabic Structures in the Ibero-Romance Lexicon

Volker Noll\*

### RESUMO

Durante o período árabe (711-1492), as línguas ibero-românicas receberam um grande número de empréstimos tanto no léxico como na toponímia e hidronímia. Como os dicionários etimológicos costumam se limitar à análise individual de cada lema, este artigo se propõe a explicar as estruturas árabes subjacentes que influenciaram nas formas ibero-românicas.

**Palavras-Chave:** empréstimos árabes, etimologia, estruturas linguísticas árabes

### ABSTRACT

Ibero-Romance integrated an important number of Arabic loanwords including toponyms and hydronyms, especially during the Moorish period (711-1492). This article aims to provide some explanatory notes on underlying Arabic language structures in these borrowings, as etymological dictionaries usually only refer to the headword in question.

**Keywords:** Arabic loanwords, etymology, Arabic language structures

*Articlista convidado*

<http://dx.doi.org/10.18364/rc.2021nEsp.509>

\*Universidade de Münster, vnoll@uni-muenster.de, orcid: <https://orcid.org/0000-0003-0261-747X>

---

1 Este artigo é a adaptação de um artigo publicado em língua alemã: “Arabische Strukturen im spanischen Wortgut”, in: FRANKE, A.-S.; ÁLVAREZ VIVES, V. (Org.). *Romaniae Pontes. Beiträge zur Sprache in der Gallo- und Iberoromania*. Berlin: Lang, 2018, p. 335-349.

## Considerações iniciais

A história de al-Ándalus começou com a conquista da Península Ibérica por tribos árabes e berberes em 711 e chegou a um final com a rendição de Granada em 1492. Do ponto de vista militar, o fim da Reconquista já se anunciou no século XIII, quando, após a unificação dos reinos de Castela e Leão, os centros do Sul foram recuperados (Córdova 1236, Sevilha 1248, Cádiz 1262). A tomada de Faro em 1249 marcou o fim da Reconquista em Portugal. A seguir, os muçulmanos que foram batizados permaneceram na Península até o começo do século XVII quando a população mourisca foi expulsada da Espanha.

O contato com a cultura árabe teve múltiplas influências na economia e nas ciências da Península Ibérica e também sobre o vocabulário das línguas autóctones. Atualmente, o dicionário da Real Academia Española contém 1.274 palavras de origem árabe (árabe dialetal, árabe clássico, mediação moçárabe e formações híbridas) (cf. DRAE<sup>23</sup>2014).<sup>2</sup> Nos dicionários atuais da língua portuguesa, o número de arabismos não alcança 700 lemas, no catalão são por volta de 350.

Num trabalho sobre o artigo *al* nas línguas ibero-românicas, Winet (2006, 299) estabeleceu, em perspectiva diacrônica, um corpus de arabismos que inclui 1.437 palavras em espanhol, 901 em português e 647 no catalão. Na área da toponímia, Lautensach (1960: 24, 32) identificou 2.343 elementos árabes na Espanha (incluindo a Catalunha). Conforme o mesmo autor, existem 290 nomes árabes ou arabizados na hidronímia da Península Ibérica, enquanto Terés (1986) lidou com até 436 hidrônimos de origem árabe em seu estudo.

---

2 No entanto, o mero número de empréstimos é relativo porque não informa sobre o uso e a frequência destes arabismos ou atribuições inseguras. Quanto a possíveis erros, o DRAE, p. ex., atribui a etimologia de *zafar* ‘livrar-se de’ a “ár. clás. *azāḥa* ‘quitar’” (2014, s.v.). Contudo, a verdadeira origem da palavra é *šāff(i)* do árabe dialetal de al-Ándalus (CORRIENTE 2008, s.v. *saḥar*). Esta solução é confirmada pela presença de *šāfi* ‘basta, chega!’ no dialeto árabe de Marrocos.

Genealógicamente, o árabe pertence às línguas afro-asiáticas (semíticas), tipologicamente é uma língua flexionada baseada em radicais consonânticos. Em regra, um radical consta de três consoantes que constituem um conceito. A partir da modificação do radical por meio de alternância vocálica, reduplicações e diversos afixos, formam-se a morfologia e o vocabulário. A fim de sistematizar o funcionamento da derivação, a gramática árabe se serve de um modelo com base na variação do radical F-Ṣ-L ‘fazer’. Deste modo, o modelo *mafʿūl* aplicado ao radical K-T-B ‘escrever’ dá *maktūb* ‘escrito; carta’.

Ao contrário das línguas europeias, a estrutura escrita do árabe, moldada com a compilação das partes até então fragmentadas do Alcorão no meio de século VII, ficou notavelmente estável ao longo dos séculos. Além da forma standardizada do árabe (árabe escrito, neo-árabe), existe um grande número de dialetos entre os quais o magrebino do noroeste da África se afastou mais do padrão do que os outros. Porém, salvo o maltês, estes dialetos não se usam na escrita exceto em transliterações destinadas à aprendizagem de idiomas. Como língua do Alcorão e vínculo dos crentes, os muçulmanos acham imprescindível preservar a unidade da língua árabe sem dar nenhuma importância oficial aos dialetos.

A forma dos arabismos nas línguas ibero-românicas depende, por um lado, do sistema da língua de chegada, sobretudo na fonética. Na morfologia, por outro lado, destacam-se também estruturas subjacentes do árabe clássico e dialetal. Como os dicionários etimológicos costumam indicar estas estruturas apenas em relação ao lema respetivo, nos propomos a analisar aqui algumas características em conjunto. Contudo, não vamos examinar paralelismos sintáticos presentes nas variedades ibero-românicas antigas. Para aprofundar aspetos da gramática árabe remetemos para a obra de Corriente (2002).

## 1. Fonética e alfabeto

### 1.1 Consonantismo e assimilação

A língua árabe possui 28 consoantes que, na gramática tradicional, são divididas em “letras solares” e “letras lunares”. Na posição inicial, as letras solares assimilam o [l] do artigo determinativo árabe *al*. A denominação particular decorre da palavra *aš-šams* ‘o sol’ que exhibe este fenômeno. Refere-se a 14 consoantes que são as interdentais /t̪/, /d̪/, /ð̪/, as dentoalveolares /t/, /d/, /t̪/, /d̪/, /n/, as alveolares /r/, /l/, /s/, /z/, /ʃ/ e a pré-palatal /ʃ/. As 14 letras lunares (< ár. *al-qamar* ‘a lua’) não assimiladoras são, na maioria, fonemas com articulação labial ou (salvo /ǧ/ [(d)ʒ]) palatal e pós-palatal /b/, /m/, /w/, /f/, /ǧ/, /j/, /k/, /ħ/ [x], /ǧ/ [ɣ], /q/, /h/, /ʕ/, /ʔ/, /h/.<sup>3</sup> Os dois tipos de aglutinação do artigo se refletem nos exemplos *alcáçova* < ár. *al-qaṣaba* e *açúcar* < ár. *as-sukkar*. Em *açúcar*, o artigo árabe foi reduzido ao *a-* inicial da palavra (cf. 2.2).

Foneticamente, o árabe não influenciou no sistema das línguas ibero-românicas. Contudo, notam-se processos adaptativos. Ao contrário das línguas românicas ocidentais, o árabe possui consoantes geminadas. É possível que a geminação tenha interferido no espanhol e no português antigos. Pois em combinação com a assimilação do artigo, a consoante longa [ss] (ár. *as-sukkar*) era interpretada como africada (esp. ant. *açúcar* [ts > s > θ]) em vez de converter-se simplesmente à sibilante [s]. Pode-se também que a associação com a africada [ts] se deva ao [s̠] apicoalveolar autóctone (que perdura no castelhano, no catalão assim como na Beira, no norte de Portugal) por ser mais chiante que o /s/ pré-dorsal árabe. Outro caso de adaptação ocorreu na preposição *ḥattā* que deu *hasta* [> st] em espanhol (pt. *até*).

No castelhano antigo, o *f-* inicial costumava transformar-se em [h-] na fala (na escrita, a partir do século XIV). Daí resulta uma interferência que incluiu também a fricativa faringal árabe [ħ]. Deste modo, *ḥattā* se anotava

3 Formalmente, deve-se notar que, na transliteração das palavras árabes, <j> é apresentada como <y>.

<fasta> ['hasta]. A fricativa velar [x], ainda ausente no espanhol antigo, foi substituída pela fricativa [f] como no exemplo *al-ḥayyāt* > esp. ant. *alfayate* (pt. *alfaiate*). Além disso, há casos de perda como em *al-maḥzan* > esp. *almacén* (cf. 2.1), pt. *armazém* (vs. cat. *magatzem*, it. *magazzino*).

Para além de velares e faringais típicas, o árabe possui quatro consoantes enfáticas (t, d, ḏ (z) s) cuja peculiaridade reside na velarização que influi também no ambiente vocálico. Nas línguas ibero-românicas, o ḏ enfático foi substituído pelo nexa [ld] (ár. *al-qāḏī* ‘juiz’ > esp., cat. *alcalde*, pt. *alcaide*; ár. *aḏ-dayʿa* ‘burgo’ > esp., cat. ant. *aldea*, pt. *aldeia*). A relação particular entre a consoante lateral e o ḏ se nota também na evolução do lat. *bellus* para *bèḏḏu* na Sicília e em partes da Calábria e da Apúlia. Não obstante, os árabes consideravam o ḏ enfático como um som que existia apenas em seu idioma e, por isso, chamavam o árabe a língua do *ḏāḏ*.

Resulta da descrição das consoantes acima que o [g] não faz parte do inventário fonético árabe (salvo em certos dialetos). Por conseguinte, o [g] românico foi substituído por [(d)ʒ] em árabe. Somente com base nesta substituição é possível explicar a evolução do lat. *Tagus* a Tejo. O espanhol que conservou a vogal original deslocou a consoante para a área velar na segunda metade do século XVI (esp. *Tajo* [ʒ > x]).

O [p] também não existe em árabe e foi substituído por [b]. Assim, Sevilha (lat. *Hispalis*) se chamava *Išbīliya* em árabe. Este topônimo ilustra também o tratamento que recebia o /s/ apicoalveolar [ʃ] castelhano em árabe que era interpretado como [ʃ], devido a sua articulação algo chiente. Consequentemente, anotava-se com <ش š> na literatura aljamiada que se servia do alfabeto árabe. Palavras espanholas alteradas por esta substituição são *jabón* (orig. germ. → lat. vulg. *\*sapone*) cuja consoante inicial velar só pode ser explicada através da prepalatal árabe [ʃ-] (cf. pt. *sabão*) assim como *jibia* (< lat. *sepia*) cuja forma antiga é preservada na variante galega *xiba*.

O nexa árabe *-st-* em *mustaʿrab* ‘arabizado’ esteve sujeito a uma metátese que resultou em *moçárabe*. Os moçárabes eram cristãos sob o domínio árabe que falavam uma variedade ibero-românica nascida durante

a latinização antiga da Bética no sul da Península. A denominação deste grupo que poderia prestar-se a confusão se deve à adoção de modos de vida árabe depois da Conquista.

## 1.2 Vocalismo e acentuação

O árabe tem um sistema de vogais de dois níveis com as vogais básicas /i/, /a/, /u/. Fonologicamente, a língua distingue também entre vogais longas e breves. Nos dialetos modernos, as condições alofônicas são diferenciadas.

Um fenômeno fonético que já existia na Idade Média é a imala (ár. *imāla*). Trata-se do fechamento das vogais tônicas /ā/, /a/ que teve como resultados [ɛ: ɛ] ou, no caso extremo, [i: i]. O segundo degrau desta evolução é chamado *imala granadina*. Destaca-se em *Bib-Rambla*, uma praça em Granada onde havia uma porta do mesmo nome (ár. *bāb* ‘porta’ > [bi:b]) e, aliás, no caso já mencionado de lat. *Hispalis* > ar. *Isbīliya* > esp. *Sevilla*. No vocabulário espanhol comum, deparamos com *albañil* ‘pedreiro’ < ár. *al-bannā*?. Elementos típicos de origem árabe na agricultura como *acequia* ‘regueira’ < ár. *as-sāqiya* e *aceña* (pt. *azinha*) < ár. *as-sāniya* mostram o primeiro degrau do fechamento. Outros elementos afetados são topônimos compostos com *Beni-* (< ár. *banū*, *banī* ‘filhos de ...’) que derivam de nomes de tribos, p. ex., *Benicasim* (Castellón) < *banī Qāsim* e *Benifató* (Alicante) < *banī Fathūn*; cf. 2.1). Atualmente, o artigo determinativo árabe *al* se pronuncia com imala até no árabe padrão ([ɛl ɛ-]). Mas os arabismos nas línguas ibero-românicas com artigo árabe aglutinado começam quase sempre com *a-* (cf. 3.2).<sup>4</sup>

Quanto à acentuação, o árabe aplica o princípio das regras do latim, ou seja, a penúltima é enfatizada se for longa, caso contrário a antepenúltima.<sup>5</sup> Enquanto o português reproduz a forma proparoxítone árabe *al-qaṣaba* com

4 Uma exceção conhecida é *elixir* (< ár. *al-ʾiksīr*). A imala não ocorre em combinação com as consoantes enfáticas (t, ḏ, ḏ, ṣ) e o [q] uvular.

5 Existem exceções como os adjetivos relacionais oxítonos em *-ī* (> esp., pt., cat. *-í*; cf. 3.5).

*alcáçova*, o espanhol adiou o acento para a penúltima (esp. *alcazaba*). Como isto ocorreu sistematicamente, formas proparoxítonas como *cómodo* se reconhecem facilmente como cultismos.

### 1.3 Alfabeto e escrita

O alfabeto árabe é perfeitamente adaptado à estrutura consonântica dos radicais de modo que é possível limitar-se à representação das consoantes e das vogais longas na escrita. Contudo, isto gera certa ambiguidade, dado nem <j> e <ī>, nem <w> e <ū> se distinguem na grafia. Existe um inventário de marcas diacríticas que indicam por cima e por baixo da linha vogais breves, a falta de vogais e consoantes longas. Mas a notação vocalizada só se usa no Alcorão e em livros didáticos. Sem vocalização, o leitor deve discernir por si mesmo se, p. ex., o carácter <j> representa <j>, <jj>, <ī>, <aj> ou <ja>.

Por conseguinte, o alfabeto árabe é deficiente na notação de línguas que têm uma estrutura morfológica diferente. As limitações são ainda mais significativas quando se trata de um vocalismo variado com três ou quatro níveis. A fim de contornar estas dificuldades, o turco adotou o alfabeto latino em 1928. Na história linguística da Península Ibérica, os problemas de notação concernem sobretudo à interpretação das carjas (esp. *jarchas*). São as estrofes finais de poemas do género chamado *muwaššah* (existe também uma série hebraica), escritas em árabe a partir do século XI que contêm também formas românicas (moçárabes). O exemplo seguinte faz parte da carja XXXV (cf. GALMÉS DE FUENTES 1994, 34):

bkʔlh ʔl ʔqd	Bokēlla <sup>h</sup> al-ʔiqdi,	Boquita de collar,
dlj km ʔlšhd	dolche komo-š-šuhdi,	dulce como la miel,
bʔn bjm	ven, bėjame.	ven, bésame.

A coluna da esquerda representa a transliteração árabe, a do meio a interpretação vocalizada e a da direita a tradução ao espanhol moderno. É

manifesto que o vocalismo românico exposto aqui se restituiu unicamente a base de formas conhecidas. Só o <-h> na primeira linha representa o -a final de *bokella* (f.), dado que a desinência do feminino árabe se nota com <-h> acrescentado de dois pontos diacríticos < ̣ > (chamado *tāʾ marbūʿa*). Por isso, a transliteração da desinência feminina árabe varia entre -a e -ah. Em <dḷj̣>, a consoante <j̣> [(d)ʒ] representa o [tʃ] românico, visto que a africada surda não faz parte do alfabeto árabe. Em <ḅj̣m> “bésame”, portanto, a grafia corresponde à fricativa sonora (cf. pt. *beija-me*).

A notação meramente consonântica dos elementos moçárabes nas carjas sugere que os escriturários tinham um bom conhecimento da variedade românica, pois, na escrita árabe, palavras estrangeiras costumam reproduzir-se com vogais longas (cf. ingl. *telephone* > ár. <tlīfūn>). Em contrapartida, na literatura aljamiada dos mouriscos batizados à força (depois de 1492), os textos espanhóis escritos no alfabeto árabe apresentam sistematicamente as marcas diacríticas. A palavra *aljamiado* deriva do ár. *ʿaġamī* ‘bárbaro, persa; não árabe’. É interessante constatar que o verbo correspondente *aġama* tem o significado ‘colocar marcas diacríticas’. O uso do alfabeto árabe para anotar textos espanhóis (sobretudo aragoneses) exprime a dedicação dos novos cristãos à sua cultura de origem.

## 2. Morfologia

### 2.1 O substantivo

O árabe distingue dos géneros gramaticais, o masculino e o feminino. Nas formas indeterminadas, o feminino termina geralmente em -a<sup>tan</sup> (*muslima<sup>tan</sup>*) que corresponde a -a [a] na posição final (transliterada -a ou -ah). O [t] da desinência só aparece na inflexão (formação do plural, construções genitivas). Na declinação, o nominativo, o genitivo e o acusativo (o dativo é preposicional) têm as desinências -u, -i, -a no masculino em nomes determinados e, respetivamente, -atu, -ati, -ata no feminino. Há

também substantivos diptóticos com nominativo e uma forma oblíqua só. Com vista às línguas ibero-românicas, a morfossintaxe árabe se manifesta sobretudo na aglutinação do artigo assim como nas construções genitivas em topônimos e hidrônimos (cf. 3.3).

Em conexão com a inflexão, é significativo que o árabe tenha reduzido e modificado as terminações na língua falada já na Idade Média. Esta forma de linguagem é comumente chamada de árabe médio (cf. VERSTEEGH 2006-09, s.v. *Middle Arabic*). A denominação se refere à reestruturação, não é entendida como uma época intermediária como, por exemplo, o francês médio. Enquanto, na língua padrão, ‘a língua árabe’ se diz *al-luġatu ‘l-‘arabīyatu*, a língua falada realiza *al-luġa ‘l-‘arabīya*.

O plural regular “sano” se forma por sufixação (*muslim<sup>un</sup>* (m.) > *muslimūn<sup>a</sup>*, *muslima<sup>un</sup>* (f.) > *muslimāt<sup>un</sup>*). Para além disso, o árabe dispõe do dual com as desinências *-ān<sup>i</sup>* (Nom.) e *-āyn<sup>i</sup>* (Gen./Ac.). Como aconteceu na formação das línguas românicas, o árabe coloquial prefere as formas flexivas oblíquas. Por conseguinte, os dialetos modernos não formam o plural em *-ūn<sup>a</sup>*, mas em *-īn* (> *muslimīn* ‘muçulmanos’). O topónimo *Alcalatén* (Castellón) representa o dual (→ ‘as duas fortalezas’) do étimo árabe *al-qal‘a* ‘a fortaleza’. Consequentemente, não se trata do nominativo *al-qal‘tān*, mas da forma oblíqua *al-qal‘tayn* como, aliás, no nome do emirado de Barém ‘dois mares’ (< ár. *baħr* ‘mar’). O caso oblíquo está também na base de topônimos de origem tribal com *Beni-* (ár. *banī*, obl. de *banū* ‘filhos de ...’). *Benicasim*, em Catalunha (cat. *Benicàssim*), vem de *banī Qāsīm*. Em contrapartida, a história da África do Norte fala dos Banū Hilāl, na forma nominativa (“-ū”) e gramaticalmente correta. Estes tribos invadiram e arabizaram o Magrebe a partir do século XI, eliminando a variedade românica ainda existente sobretudo na área da Tunísia.

O árabe forma *nomina loci* a base de participios substantivados, prefixando, p. ex., o morfema *ma-* em *maħzan* < *ħazana* ‘guardar’ que deu *armazém* em português. *Magrebe* (< ár. *maġrib* ‘pôr-do-sol’) contém o participio ativo do verbo *ġaraba* ‘pôr-se’ (modelo *maf‘il*). Outra palavra

baseada no mesmo radical é *ġarb* ‘oeste’ que, com o artigo aglutinado, deu o nome ao Algarve. *Almocadém* (esp. ant. *almocadán*) provém do participio passivo substantivado *al-muqaddam* < ár. *qaddama* ‘antepor’ formado a partir do radical verbal II (cf. 2.4) e caracterizado pelo alongamento da segunda consoante (modelo *mufa<sup>sc</sup>al*).

O diminutivo árabe se forma segundo o modelo *fu<sup>sc</sup>ayl*. Assim, *albufeira* (esp. *albufera*) deriva do ár. *buhayra* ‘lago’ < *bahr* ‘mar’. Como a fricativa faringal surda [ħ] não existe nas línguas ibero-românicas, foi substituída pela fricativa [f] (cf. 2.1).

## 2.2 A aglutinação do artigo *al*

A aglutinação do artigo é sem dúvida a particularidade mais notória dos arabismos nas línguas ibero-românicas. Ao contrário do artigo determinativo românico, o artigo árabe *al* não indica nem gênero, nem número. Desde o início do século XX, os linguistas estavam conscientes de que havia um contraste entre o espanhol e o italiano quanto à aglutinação. Em 1931, o romanista v. Wartburg opôs arabismos com e sem artigo aglutinado em espanhol, português e catalão às formas geralmente sem artigo nos arabismos do italiano (1931, 226-227). A aglutinação ocorre também nas variedades do berbere, enquanto os arabismos nas outras línguas (turco, persa, suaíli, malaio etc.) carecem de formas aglutinadas.

Graças ao estudo de Winet (2006, 298-299; 319), é possível quantificar diacronicamente a extensão da aglutinação nas línguas ibero-românicas. O número de formas aglutinadas em relação às formas sem artigo é de 855 : 840 em espanhol, 517 : 518 em português e 309 : 434 em catalão. Há também palavras que apresentam ambas as variantes no decorrer da sua história (258 em espanhol, 134 em português e 96 no catalão).

Conforme estes resultados, o espanhol conta com um pouco mais de 50% de formas aglutinadas, enquanto, no catalão, a formas sem artigo montam a 70% (cf. esp. *algodón* vs. cat. *cotó*). A presença ou a falta do artigo

árabe (*al*, *aC-*) pode até indicar o caminho que certos arabismos tomaram antes de chegar ao centro da Europa, dado que as rotas principais originavam na Espanha e no sul da Itália. Assim, o contraste que existe entre fr. *sucre* e alem. *Zucker* frente a esp. *azúcar* aponta para a transferência italiana (cf. it. *zucchero*, lat. medieval *zuccarum*).

Desde os anos 1930, a linguística tentou explicar por qual motivo arabismos apresentavam o artigo aglutinado ou não (cf. o relatório em Noll 1996, 2006). A teoria preferida tem sido a da influência berbere (cf. CORRIENTE 2003, 61). Segundo esta teoria, é o emprego abusivo do artigo na fala árabe dos berberes que teria causado a aglutinação. Na verdade, isto é uma alegação que não explica nem o mecanismo da aglutinação, nem o paralelismo de formas com e sem artigo, nem a situação algo diferente no catalão. Não se sabe muito da fala árabe dos berberes na Península Ibérica. Contudo, não é provável que ficasse gramaticalmente defectiva através dos séculos, já que, em meados do século IX, eclesiásticos lamentavam que os moçárabes se exprimissem melhor em árabe que em latim. É preciso dar uma explicação que tome mais em consideração os próprios fatores linguísticos.

O artigo árabe *al* tem uma vogal inicial instável (ár. *alif al-wašl*). Quando as desinências flexionais (*-u*, *-i*, *-a*) deparam com o artigo, elidem sintaticamente sua vogal inicial. O contato com as 14 letras lunares (cf. 1.1) que não assimilam o [l] do artigo leva a 42 combinações possíveis do jeito { [ul il al]+letra lunar inicial }. 14 delas têm a forma *al*. O contato com as 14 letras solares gera igualmente 42 combinações. Mas dada a assimilação regressiva junto com a elisão da vogal inicial do artigo, resultam 42 alomorfos como {*a\_ 's*}-*s*, {*u\_ 't*}-*t* etc. Por conseguinte, a forma *al* só ocorre 14 vezes entre as 84 combinações possíveis. É claro que este cálculo nem tem em conta a frequência das formas, nem a redução das desinências na língua falada. Contudo, é óbvio que a grande variabilidade sintática do artigo contribuiu a impedir a aglutinação na maioria das línguas que integraram arabismos.

Acresce que intervém outro fenômeno fonético-sintático que resulta da elisão da vogal inicial do artigo. A perda da vogal (*al* > *'l*; *'s* etc.) que iguala

à perda de uma sílaba faz com que o artigo restante (p. ex. [l]) seja enclítico ao elemento que precede e, com isto, separado silabicamente do substantivo que determina (→ ár. *šufti\_ʾl* | *-walad* ‘você viu o menino’). Na línguas crioulas de base francesa, há centenas de exemplos que mostram, que a divisão silábica interferiu com a segmentação morfemática, incorporando uma parte do artigo ao substantivo. P. ex., fr. *les oiseaux* [le-zwazo] ‘os pássaros’ – parece que o plural da palavra é mais frequente que o singular – tornou-se *zwaso* ‘pássaro’ em crioulo (cf. BAKER 1984).

Na variedade árabe de al-Ándalus (esp. *árabe andalusí*), os dados linguísticos apresentam uma particularidade. Corriente apontou para o fato que, ao contrário das outras variedades, o árabe de al-Ándalus não praticava a elisão da vogal inicial do artigo, mantendo *a-* em todas as circunstâncias.<sup>6</sup> No contexto das explicações dadas acima, esta particularidade anula a separação silábica entre o artigo e o substantivo que determina (→ ár. *šufti* | *al-walad*). Além disso, a vogal inicial *a-* é conservada em todas as 84 combinações sintáticas consideradas e, na metade dos casos (→ 42) quando se trata de letras lunares, a forma é mesmo *al*. Devido à frequência, *a-*, *al* devem ter adquirido quase um status de sinal identificador para substantivos.

O paralelismo de formas aglutinadas e de empréstimos sem artigo nas línguas ibero-românicas (e as variedades berberes) tem uma explicação sociolinguística. Al-Ándalus e o Noroeste da África são territórios onde existia um bilinguismo prolongado com a língua árabe. Isto nunca foi o caso

6 “D’une façon tout a fait exceptionnelle, l’article défini en arabe andalou *ʔal+* ne perd pas sa voyelle sous l’élision, comme cela se produit dans le reste des formes de l’arabe ancien, arabe standard et néo-arabe, ce qui se reflète dans les graphies arabes comme dans IQ 9/4/4 >tağli alrāḥah< « tu retabliras le repos », 40/4/2 >alṣaqlī arrāğih< « la raison pondérée », et dans les transcriptions romanes comme dans Alcalá 9 *naāti al kóbz* « je donne le pain », *mi alhayé* « eau-de-vie », mais aussi dans quelques noms de lieu : *Benialfaquí* (Alicante) < arabe andalou *bani ʔalfaquí* « les fils du fakih », *Binialmara* (Mallorca) < arabe andalou *bani ʔalmará* « les fils de la femme ». Même avec les prépositions *bi+*, *li+* et *fi+* on trouve des contractions avec chute de la voyelle /i/ [...]” (CORRIENTE et al. 2015, 103).

nas sociedades turca, persa etc. Na Sicília, a dominação árabe (827-1091) não perdurou, limitando-se a zona de influência de Palermo.

Por conseguinte, as formas ibero-românicas aglutinadas podem ser explicadas pela influência de dois fatores significativos. Estes são os meios bilíngues dos moçárabes em al-Ándalus em conexão com as partículas iniciais constantes *a-*, *al-* facilmente reconhecíveis. O menor número de aglutinações em catalão se conforma diacronicamente com a Reconquista antecipada do centro da Catalunha, o condado de Barcelona (801). As formas sem artigo têm origem no meio de contato tradicional, que naturalmente também existia na Península Ibérica como na Itália e em todas as outras áreas de contato no Oriente.

Há também empréstimos com o artigo aglutinado que foram introduzidos pela escrita, ou seja, pelas traduções. Neste contexto é importante destacar que o artigo árabe sempre forma uma unidade gráfica (e ótica) com o substantivo que segue. A despeito da assimilação fonética, a escrita nota por princípio ال *al* (p. ex., ريح <rīḥ> ‘vento’ → الريح <al-rīḥ> [a’rī:h] ‘o vento’). A adoção irrefletida da forma escritural se reconhece às vezes quando aparecem formas erradas que estão em desacordo com a realização fonética. Nos nomes de estrelas, nota-se a falta de assimilação de [l] diante as letras solares em *Aldebarã* (< ár. *ad-dabarān* ‘o que segue [as Plêiades]’) (constelação Touro) e em *Altair* (< ár. *aṭ-ṭā’ir* ‘o pássaro’) (constelação Águia). Neste último caso, existe também a forma correta *Atair*.

A proeminência do artigo árabe se repercute também em formas misturadas ou híbridas. Em *almirante* (cat. *almirall*) que deriva do ár. *ʿamīr* ‘príncipe, comandante’, *al-* não faz parte do étimo. Há até formas híbridas com elementos românicos como, no espanhol, *almatriche* ‘regueira’ (< ár. *al* + lat. *matrix*), *almena* ‘merlão’ (< ár. *al* + lat. *minae*), *azufre* ‘enxofre’ (*a* + lat. *sulphur*) (cf. MÜLLER 2004) e o topónimo *Alpuente* (Valencia). Estas formações realçam o papel de *al-* como sinal identificador que serviu até para estabelecer uma linhagem árabe de prestígio na linguagem.

### 2.3 Topônimos e hidrônimos. O *status constructus*

Na hidronímia, a Península Ibérica registra numerosos nomes de origem árabe que têm uma conexão com *Guad-* (em Portugal também *Od-*). Esta é uma peculiaridade linguística na medida em que os nomes dos rios estão geralmente entre os elementos mais estáveis diacronicamente que existem desde os tempos mais remotos. *Guad-* deriva do ár. *wād(in)* (ou *wādī* + N). Ao contrário do ár. *nahr* ‘rio’, denota um leito de rio, o que reflete as condições climáticas do sul da Europa e do Oriente, onde os rios podem secar na estação quente.

O desenvolvimento fonético da palavra segue uma tendência românica de evitar a labiovelar [w] em conexão com [a] no nexa inicial *wa-* e de apoiá-lo por uma consonância protética. O resultado foi o nexa *gua-* com a velar homorgânica protética [g]. O mesmo desenvolvimento ocorreu nos empréstimos germânicos respetivos (cf. germ. *\*wardōn* > it. *guardare*, esp. *guardar*, fr. *garder*).

Os topônimos e hidrônimos constituem com frequência grupos nominais cuja relação se define, conforme a gramática árabe, como *status constructus*. Nesta construção, ambos os substantivos estão determinados, embora o artigo só apareça diante do genitivo que segue. *Guadalajara* deriva da forma árabe dialetal *wād al-ḥaġāra* (ár. *wādī 'l-ḥiġāra*) ‘o álveo das pedras’. *Guadiana* é uma formação híbrida com um nome próprio (< ár. *wādī* + *Ana*). É possível que tenha conservado a vogal final *-ī* de *wādī*, portanto, existem também variantes do nome como *Yāna* (cf. TERÉS 1986, 454).

O exemplo do rio Guadiana mostra que, no caso de nomes próprios, o artigo árabe é omitido. Trata-se de uma justaposição como em *Calatayud* (Zaragoza) < ár. *qal'at 'Ayyūb* ‘a fortaleza de 'Ayyūb’ (Jó) e também *Gibraltar* < ár. *ġabal Ṭāriq* ‘o monte de Tárique’.

Quando um adjetivo atributivo e um substantivo se combinam, ambos levam o artigo em árabe, como é o caso no superlativo francês (*la maison la plus riche* ‘a casa mais rica’). Uma exceção notável é *Guadalquivir* (< ár.

dialetal *al-wād al-kibīr*, ár. *al-wādī 'l-kabīr* ‘o rio grande’). A omissão do artigo antes do substantivo aqui só pode ser explicada em analogia com as numerosas formas em *Guad-*.

## 2.4 O verbo

Chama a atenção o fato que, entre os empréstimos árabes, se encontram poucos verbos. O dicionário da Real Academia Española (DRAE 2014) conta apenas 30 que, na maioria, são derivações denominais. Isto está sem dúvida relacionado à estrutura dos radicais árabes a qual implica uma alta variabilidade morfológica dos verbos. Como não há infinitivo, a forma canônica do verbo é a 3ª pessoa do singular do pretérito perfeito (*kataba* ‘ele escreveu’). Neste caso, as especificações número-pessoais se realizam por meio de sufixos (*katabta* ‘você escreveu’), o aspeto imperfectivo é formado tanto com prefixos como sufixos (*yaktubūn*<sup>a</sup> ‘eles escrevem’). Além disso, existe um sistema de derivação com dez formas verbais (listadas com os numerais romanos I-X) que acarretam nuanças de sentido. A forma VIII, p. ex., introduz o infixo {t} no radical K-T-B: *kataba* ‘escrever’ > *iktataba* ‘firmar’, *istiktāb* ‘ditado’. Em comparação com as bases facilmente reconhecíveis das línguas românicas, este sistema é bastante opaco quando se trata de identificar uma forma verbal própria para o empréstimo.

Não obstante, a forma IV do verbo árabe, com prefixação de *a-* e significado factitivo (ár. *kataba* > *aktaba* ‘ditar’), foi colocado em conexão causal com formações verbais parecidas nas línguas ibero-românicas (cf. SALOMONSKI 1944). Todavia, não é preciso recorrer ao modelo árabe para explicar uma forma como *assentar*, dado que existe uma alternativa morfológica evidente que é a prefixação latina com *ad-* (*\*adsedentare* < lat. *sedēre*). Até no caso de *aforrar* (esp. ant. *aforrar*, esp. *ahorrar*) cujo étimo é mesmo árabe, a referência à forma IV não é contundente. Pois na língua árabe, só existe a forma II deste verbo (ár. *ḥarrara* ‘liberar’). Contudo, outra palavra derivada do mesmo radical é o adjetivo *ḥurr* ‘livre’ que deu *forro*,

*horro* ‘livre’ em português, catalão e espanhol (cf. pt. *carta de alforria*). No caso dos verbos *aforrar*, *ahorrar* se trata de uma derivação deadjetival, visto que, na forma catalã *alforrar*, destaca-se a parte <a> que é o artigo árabe. Por conseguinte, o étimo que deu origem à derivação deve ser o adjetivo substantivado *al-ḥurr* ‘o [escravo] libertado’ (→ *alforrar*, *ahorrar*).

## 2.5 O adjetivo

Na língua espanhola, há por volta de 130 adjetivos de origem árabe (cf. DRAE 2014). 33 destes empréstimos têm uma forma particular oxítone em *-í* (pl. *-ies*) e carecem da terminação específica do feminino. Derivam do adjetivo relacional árabe, chamado *nisba*, com as desinências *-ī* (m.) e *-īya* (f.) em árabe (p. ex., ár. *‘arabī*, *‘arabīya* ‘árabe’). Estas formações se referem ao mundo árabe-islâmico (esp. *marroquí*, *alfonsí*), mas pararam de ser produtivas nas línguas ibero-românicas. O português e o catalão têm um número reduzido das formações árabes em /i/ final (cf. esp. *marroquí* vs. pt. *marroquino*, *-a* vs. cat. *marroquí*, *-ina*; esp. *paquistaní* vs. pt. *paquistânês*). Existem também nominalizações como *javali* (esp. *jabali*) < ár. *ǧabalī* ‘montanhoso, do monte’ e *maravedi* (< ár. *murābiṭī*), uma moeda cunhada pelos almorávides na Península Ibérica.

No nome da Alhambra ‘a (fortaleza) vermelha’, deparamos com outra formação adjetival. Em árabe, os adjetivos de cor têm formas peculiares. *Ḥamrāʾ* (modelo *faʿlāʾ*) é o feminino de *ʾaḥmar* ‘vermelho’. Conforme a fonotática da língua espanhola e as leis de contato de sílabas, o nexos heterossilábico *-m|r-* integrou a consoante epentética [b] a fim de reduzir a sonoridade do ataque da segunda sílaba. Outra formação adjetival em combinação com uma cor é Canilla de Albaida, um município na província de Málaga. Como Alhambra, *Albaida* junta o artigo determinativo a um adjetivo de cor árabe no feminino. A forma *bayḍāʾ* ‘branca’ (m. *abyaḍ*) é a mesma que faz parte do nome autóctone da cidade marroquina de Casablanca (ár. *ad-dār al-bayḍāʾ*).

## Bibliografia

BAKER, Ph. The Significance of Agglutinated French Articles in the Creole Languages of the Indian Ocean and Elsewhere. In: SEBBA, M; TODD, L. (Orgs.). **Papers From the York Creole Conference**, 24–27 September 1983. University of York, Department of Language, p. 1-10, 1984.

CORRIENTE, F. **Gramática árabe**. Barcelona: Herder, <sup>2</sup>2002.

CORRIENTE, F. **Diccionario de arabismos y voces afines en iberorromance**. Madrid: Gredos, <sup>2</sup>2003.

CORRIENTE, F. **Dictionary of Arabic and Allied Loanwords. Spanish, Portuguese, Catalan, Galician and Kindred Dialects..** Leiden – Boston: Brill, 2008.

CORRIENTE, F.; PEREIRA, Ch.; VICENTE, Á. **Aperçu grammatical du faisceau dialectal arabe andalou. Perspectives synchroniques, diachroniques et panchroniques**. Berlin – Boston: de Gruyter, 2015.

DRAE. Real Academia Española. **Diccionario de la lengua española**. Madrid: Gredos, <sup>23</sup>2014.

LAUTENSACH, H. **Maurische Züge im geographischen Bild der Iberischen Halbinsel**. Bonn: Dümmlers, 1960.

MÜLLER, B. Die Arabisierung romanischer Wörter im Spanischen des Mittelalters. In: GIL, A.; OSTHUS, D.; POLZIN-HAUMANN, C. (Orgs.). **Romanische Sprachwissenschaft**. Zeugnisse für Vielfalt und Profil eines Faches. Festschrift für Christian Schmitt zum 60. Geburtstag. I. Frankfurt/M.: Lang, p. 203-211, 2004.

NOLL, V. Der arabische Artikel *al* und das Iberoromanische. In: LÜDTKE, Jens (Org.). **Romania Arabica. Festschrift für Reinhold Kontzi zum 70. Geburtstag**. Tübingen: Narr, p. 299-313, 1996.

NOLL, V. La aglutinación del artículo árabe *al* en el léxico español. In: ARNOLD, Rafael; LANGENBACHER-LIEBGOTT, J. (Orgs.). **Cosmos Léxico. Contribuciones a la Lexicografía y a la Lexicología Hispánicas**. Frankfurt/M.: Lang, p. 35-49, 2006.

NOLL, V. Arabische Strukturen im spanischen Wortgut. In: FRANKE, A.-S.; ÁLVAREZ VIVES, V. (Orgs.). **Romaniae Pontes. Beiträge zur Sprache in der Gallo- und Iberoromania**. Berlin: Lang, p. 335-349, 2018.

SALOMONSKI, E. **Funciones formativas del prefijo *a* estudiadas en el castellano antiguo**. Diss. Zürich, 1944.

TERÉS, E. **Materiales para el estudio de la toponimia hispanoárabe**. Nomina fluvial. Madrid: CSIC – Instituto de Filología, 1986.

VERSTEEGH, K. (Org.). **Encyclopedia of Arabic Language and Linguistics**. 5 vol. Leiden: Brill, 2006-09.

WARTBURG, W. v. Grundfragen der etymologischen Forschung. **Neue Jahrbücher für Wissenschaft und Jugendbildung** 7, p. 222-235, 1931.

WINET, M. **El artículo árabe en las lenguas iberorrománicas** (aspectos fonéticos, morfológicos y semánticos de la transferencia léxica). Córdoba: Universidad de Córdoba, 2006.