

O léxico português hereditário à luz da etimologia românica: reflexões a partir do DÉRom (*Dictionnaire Étymologique Roman*)

Myriam Benarroch
Université Paris-Sorbonne

Abstract

How can the comparative method applied to Romance Languages yield interesting results for Portuguese-inherited etymology is the issue which is addressed in the present paper. After presenting the project called 'DÉRom' (*Dictionnaire Étymologique Roman*), which is a new Romance etymological dictionary based on common ancestor's reconstruction from Romance cognates, we will show, from a set of articles of this dictionary, how the methodology employed here can give more specific indications of the etymons of Portuguese lexemes, delineate the history of their phonetic and morphological evolution and of their semantic variation and improve on the dating so far suggested by Portuguese and Brazilian dictionaries.

1. Introdução

A etimologia românica beneficiou, ao longo do século passado, de uma série de avanços teóricos e metodológicos, e também tecnológicos, consideráveis¹. O recurso à fonética histórica, à dialectologia, à geografia linguística e à sociolinguística permitiu a publicação de dicionários etimológicos de excelência, tais como o FEW para o francês e, mais recentemente, o LEI para o italiano. A integração destes princípios e métodos e a sua síntese na elaboração dos referidos dicionários conduziram, do ponto de vista teórico, ao que Chambon (no prelo) chama de *paradigma wartburgiano*, simbolizado pelos nomes de von Wartburg, Baldinger e Pfister. À concepção tradicional da etimologia apenas como procura da origem da palavra foi completada pelo que Baldinger chamou de *etimologia-história da palavra*, com forte preocupação semântica, uma espécie de biografia da palavra de que “o nascimento, de que se ocupava exclusivamente a etimologia antiga, não é senão o ponto de partida” (Baldinger, 1959: 239).

Do lado da etimologia portuguesa, os estudos sobre o léxico português medieval desenvolveram-se sensivelmente no último quarto do século passado (*cf.* Maia, 1986; Mattos e Silva, 2008) e conheceram um arranque notável na última década tanto em Portugal (*cf.* Castro, 2004; 2008; Emiliano, 2003a; 2003b; Martins, 1999; 2007) como na

¹ Agradecemos imenso a releitura atenta deste trabalho por Maria Alice Fernandes. *Textos Seleccionados, XXVIII Encontro Nacional da Associação Portuguesa de Linguística, Coimbra, APL, 2013, pp. 149-168, ISBN: 978-989-97440-2-8*

Galiza (cf. Boullón, 2007; Souto Cabo, 2006; e as obras informatizadas: DDGM; TMILG). Estão em curso vários projetos na área da dialetologia e da atlantologia, na edição e digitalização de textos medievais e na reflexão sobre a história da língua, tanto em Portugal como no Brasil. O conjunto destes projetos permitirá a realização de um muito esperado dicionário histórico do português. Por enquanto, não existe um dicionário etimológico da língua portuguesa que incorpore os avanços metodológicos da etimologia românica e recolha o fruto dos estudos recentes.

O ponto de partida dos dicionários etimológicos modernos da língua portuguesa, o dicionário de Antenor Nascentes, publicado em 1932 e reeditado em 1955, prefaciado por Meyer-Lübke, é de uma grande qualidade científica e foi objeto de várias publicações (cf. Lisboa, 1937; Piel, 1932); no entanto, ele não integra a data da primeira atestação nem as abonações históricas dos lexemas estudados. O DELP do saudoso José Pedro Machado, publicado por primeira vez em 1952, continua a ser, na sua terceira edição de 1977, a obra etimológica de referência: é até agora, o único dicionário a dar não apenas o étimo da palavra (às vezes com explicação etimológica), mas também a data de primeira documentação e as abonações históricas; a terceira edição, porém, não aproveitou as numerosas observações e correções formuladas por Ramón Lorenzo (1968). Os trabalhos de Antônio Geraldo da Cunha são certamente os que mais renovaram a etimologia do português nas últimas décadas. O seu dicionário etimológico (DENF), publicado em 1982, apesar de sintético, traz informações históricas, fornece a data da primeira atestação, mas não a abonação correspondente. Este dicionário foi completado pelos três fascículos do *Índice do Vocabulário do português medieval* (Cunha, 1986-1994), que abrangem as letras A-D, com um suplemento (Cunha, 1992). As abonações completas, referenciadas, foram reunidas no *Vocabulário Histórico-Cronológico do Português Medieval*, publicado em 2002 sob forma de DVD infelizmente após a morte do autor (falecido em 1999) e reeditado em 2006-2007)². Os trabalhos etimológico-históricos de Cunha foram abundantemente utilizados pelo dicionário *Houaiss*, que também aproveitou os materiais do DELP₃. Apesar de não ser um dicionário etimológico propriamente dito, o *Houaiss* é hoje em dia, entre os dicionários portugueses, o que comporta a matéria etimológica mais atualizada. No entanto, o *Houaiss* não aproveitou todo o trabalho do DELP₃, nem todas as informações que figuram nas publicações de Cunha (cf. Benarroch, 2010; Monjour, 2004). Além da lexicografia portuguesa, encontram-se informações sobre a etimologia do léxico português

² Entre as obras etimológicas de Antônio Geraldo da Cunha encontram-se também *Os estrangeirismos da língua portuguesa* (2003, São Paulo: Humanitas) e o *Dicionário histórico das palavras portuguesa de origem tupi* (1982 [1978]¹, São Paulo: Melhoramentos).

em dicionários etimológicos de outras línguas, como o DCECH, o FEW e o LEI. Quem quiser hoje fazer etimologia do português tem portanto de reunir todo esse material, e também, é claro, de andar sempre à procura de novas fontes. É o que começámos a fazer no *Dictionnaire Étymologique Roman* (DÉRom). Tentaremos mostrar aqui, na continuação de um artigo precedente (Benarroch, no prelo 1), como tanto a metodologia desenvolvida neste dicionário como as fontes utilizadas podem contribuir para melhorar a etimologia da língua portuguesa.

2. Apresentação do DÉRom

O DÉRom (*Dictionnaire Étymologique Roman*) é fruto de um projeto franco-germânico que foi inaugurado no XXV^o Congresso Internacional de Linguística e Filologia Românicas, em Innsbruck, em 2007 (*cf.* Buchi/Schweickard, 2010). O dicionário começou a ser publicado em 2008 sob forma eletrónica e está em acesso livre na internet. A equipa do DÉRom, dirigida por Éva Buchi e Wolfgang Schweickard, é composta por 71 membros, oriundos de 12 países, entre os quais redatores e revisores, duas documentalistas e um técnico de informática. O DÉRom apresenta-se como um herdeiro, embora parcial, do *Romanisches Etymologisches Wörterbuch* (REW₃) de Meyer-Lübke (Buchi/Schweickard, 2009). Parcial, porque o objetivo a curto prazo é tratar uma nomenclatura seletiva, constituída pelas 488 bases etimológicas recenseadas por Fischer (1969), comuns na totalidade ou em grande parte ao conjunto das línguas românicas

2.1. Os fundamentos teóricos e metodológicos do DÉRom

A principal inovação do DÉRom em relação à etimologia românica assenta na aplicação sistemática do método da gramática comparada – reconstrução da matéria românica. Este método, reconhecido e abundantemente utilizado para as outras famílias linguísticas (índo-europeia, bantu, austronésia e ainda dentro da família indo-europeia, nos ramos germânico, eslavo e céltico), não o foi para as línguas românicas, com o argumento de que conhecemos a língua-mãe dessas línguas – o latim, que beneficia de abundante documentação. As duas publicações de Jean-Pierre Chambon (2007; 2010) que advogam a aplicação da reconstrução comparativa às línguas românicas constituem os fundamentos teóricos em que se apoia a metodologia desenvolvida no DÉRom (*cf.* também Dardel, 2007; 2009). Trata-se de substituir o procedimento prospectivo tradicional descendente – do latim (clássico) para as línguas românicas – por um procedimento retrospectivo ascendente – dos falares românicos ao latim –, que é o da gramática comparada –

reconstrução. (cf. Buchi, no prelo; Schweickardt, 2012). Isso significa que o ponto de partida da análise etimológica não são os lexemas que se podem encontrar nos dicionários latinos e sim os cognatos românicos que permitem reconstruir o antepassado comum, a protolíngua, no sentido de Campbell (2004: 125) aqui, os étimos protorromances. Estes étimos são os verdadeiros étimos no sentido da *etimologia proxima* (cf. Meier 1964: 105) já que é do latim falado que provêm os lexemas românicos e não do latim escrito da Antiguidade que, no melhor dos casos, remeteria para a *etimologia remota*.

As normas de redação, indispensáveis para a homogeneidade dos artigos, refletem as opções metodológicas ao mesmo tempo que integram alterações, sempre que aparece uma nova dificuldade que o esquema precedente não permitia resolver. Essas normas encontram-se agrupadas num opúsculo intitulado *Livre bleu* (LB) que está permanentemente atualizado na sua versão em linha. A bibliografia utilizada para a redação dos artigos divide-se em duas categorias: uma “Bibliografia de consulta e de citação obrigatórias” (LB: 215) e uma “Bibliografia geral”. A primeira comporta 130 títulos, que devem ser consultados imperativamente para a redação de cada artigo. A segunda, que hoje consta de 1237 títulos, reúne as fontes utilizadas na redação e na revisão dos artigos.

2.2. A estrutura do artigo

Encontram-se atualmente publicados em linha 70 artigos do DÉRom e uma centena estão em fase final de redação. As entradas do dicionário são constituídas pelos étimos protorromances, reconstruídos segundo o método da gramática comparada – reconstrução a partir dos cognatos românicos. Os étimos apresentam-se sob a forma fonológica e estão precedidos por um asterisco (ex. “*/ka'ten-a/”) que significa que o étimo foi reconstruído. Esta apresentação não dificulta a consulta do dicionário: todos os artigos podem também ser interrogados a partir dos correlatos latinos (*catena*) ou das entradas do REW (*catēna*), ou ainda dos continuadores românicos (dacoromeno *cătină*, friulano *ciadène*, português *cadeia*, por exemplo). Um artigo do DÉRom comporta obrigatoriamente as partes seguintes:

(1) *Lema etimológico*: significante, categoria gramatical e significado do étimo protorromance.

(2) *Materiais*: contém o conjunto dos cognatos românicos que permitem reconstruir o étimo protorromance; são considerados vinte idiomas – não apenas línguas standardizadas mas também dialetos, essenciais para a reconstrução (cf. Andronache, no prelo) –, de entre

os quais são citados obrigatoriamente os que comportam continuadores do étimo³; para cada idioma, figuram a abreviatura do glotónimo, o significante, a categoria gramatical, o significado, a data da primeira atestação, as referências bibliográficas e a forma da variante atestada quando difere da do significante.

(3) *Comentário*: “explicita a análise dos dados reunidos na seção dedicada aos materiais que leva a propor o étimo citado na entrada do artigo [...]” (LB: 58)⁴.

(4) *Bibliografia*: “cita, por ordem cronológica, as publicações [...] que apresentam um interesse geral (*i.e.* mais ou menos panromânico) para o artigo [...]” (LB: 62).

(5) *Assinaturas*: “apresenta o conjunto das pessoas que contribuíram para a elaboração [do artigo], façam ou não parte do projeto DÉRom” (LB: 63).

(6) *Data de publicação em linha deste artigo*: com duas datas, a da primeira publicação em linha e a da última alteração efetuada no artigo.

(7) *Notas*: reúnem observações demasiado longas ou específicas para serem inseridas no texto.

2.3. Galego e português no DÉRom

Desde a redação do primeiro artigo, foi posta a questão do glotónimo que devia ser utilizado para designar o galego e o português medievais, sabendo que grande parte das fontes são as mesmas dos dois lados da fronteira. Foi objeto de numerosas discussões. As soluções adotadas evoluíram à medida que foram aparecendo novos artigos e novos problemas. A solução atual – não isenta de problemas – considera três casos:

(1) Quando um lexema comum ao galego e ao português está documentado antes de meados do século 14, leva o glotónimo “**gal./port.**”, seguido do lexema. Ex.: “**gal./port. ponte**”.

(2) Se os significantes do cognato galego e português são distintos, as informações apresentam-se sob a forma: ““**gal.**” + cognato galego + “/” + “**port.**” + cognato português. Ex.: “**gal. febreiro/port. fevereiro**”.

(3) Se o galego e o português tiveram na Idade Média um representante hereditário comum que foi eliminado em galego ou em português contemporâneo em proveito de um lexema não hereditário, então ambos os idiomas passam a ser distinguidos sob os

³ Damos aqui os glotónimos na ordem em que aparecem citados no DÉRom, de leste para oeste, com as suas abreviaturas (adaptadas ao português): dacorromeno (dacorrom.), istrorromeno (istrorrom.), meglenorromeno (meglenorrom.), arromeno (arrom.), dálmata (dál.), istriota (istriot.), italiano (it.), sardo (sard.), friulano (friul.), ladino (lad.), romanche (romanch.), francoprovençal (frpr.), francês (fr.), occitano (occit.), gascão (gasc.), catalão (cat.), asturiano (ast.), espanhol (esp.) galego (gal.), português (port.). A vogal *a-* que precede qualquer um desses glotónimos significa “antigo”, ex. aport. = português antigo.

⁴ A tradução em português é nossa.

respetivos glotónimos “gal.” e “aport.” (português antigo) ou “agal.” (galego antigo) e “port.” (cf. LB: 51).

O português está representado em 64 dos 70 artigos publicados. Desses 64 artigos, 32 lexemas têm um significante português diferente do do galego. As diferenças podem ser gráficas (gal. *allo*/port. *alho*) ou fonéticas (*fuxir/fugir*; *fabalfava*), refletindo as respetivas evoluções fónicas na diacronia das duas línguas. Algumas delas revelam o carácter mais antigo da forma galega contemporânea e a inovação que representa a portuguesa: *fame/fome*; *hedralhera*; *herbalerva*. Apenas um lexema não tem cognato em galego: é o substantivo português *louro* “loureiro”. Alguns lexemas existiram em galego antigo e/ou português antigo e desapareceram da língua contemporânea, em que venceu um concorrente, seja de ambos os lados da fronteira (*ascoitar*, com os concorrentes gal. *escoitar* e port. *escutar*; *exir*, eliminado pela concorrência de gal. *sair*, port. *sair*; *leixar*, eliminado por *deixar*; *logo*, s.m., por *lugar*), seja apenas de um só (gal. *caer*, mas port. *cair* por mudança flexional; gal. *feo*, port. ant. *fêo*, mas port. *feno* por restauro da forma culta).

3. Em que medida o DÉRom beneficia a etimologia portuguesa ?

3.1. A questão do étimo próximo : o latim falado reconstruído

Os dicionários etimológicos do português raramente dão como étimo uma forma que não seja a do latim escrito da Antiguidade (ou latim “clássico”). No entanto, muitas vezes observamos discrepâncias entre esta forma do latim clássico e o lexema português supostamente oriundo dela. E é lógico, já que é da língua falada que nasceram os idiomas românicos e não do latim escrito. A reconstrução românica permite ter acesso, embora parcialmente, ao latim falado. Os étimos protorromances reconstruídos, que constituem os significantes dos lemas do DÉRom, refletem mais fielmente os lexemas deste latim falado (cf. Benarroch, no prelo 2; no prelo 3; Buchi 2010).

Tomemos o exemplo de port. *escutar*. No DELP₃ como no Houaiss, o étimo proposto é “lat[im] *auscultāre*”. Pelo contrário, no DÉRom, a reconstrução do étimo protorromance, ao comparar todos os cognatos românicos, mostra que as formas *ascoitar*, *ascuitar*, de um lado e *escoitar*, *escuitar* de outro, remontam a duas bases etimológicas distintas, */as'kult-a-/ e */es'kult-a-/, o que explica que o português medieval tivesse duas séries, com mudança da primeira vogal: *ascoitar*, *ascuitar* (séc. 13) e *escoitar*, *escuitar* (respetivamente séc. 13 e 14). O verbo */es'kult-a-/ provém do primeiro não por uma

evolução fonética hereditária, mas sim por substituição do que Chauveau (FEW 25: 1060a) considera como uma *greffe prefixale*. No caso de duas bases etimológicas irreduzíveis, o DÉRom considera que cada base representa o lema de um artigo distinto: assim temos dois artigos */as'kult-a-/ e */es'kult-a-/ (ambos, Schmidt, in DÉRom).

Temos também o caso de duas variantes de uma mesma base etimológica irreduzíveis a um étimo único. O lema comporta então um duplo significante protorromance, e o artigo, duas subdivisões que correspondem aos dois protótipos. É o exemplo do artigo */ϕen-u/ ~ */ϕen-u/ (> port. ant. *fẽo*).

Um caso muito frequente é o do artigo em que a reconstrução românica, apesar de remontar a uma base única comum a todos os idiomas, revela fases intermediárias, cada fase tendo o seu próprio étimo próximo. Aqui, o artigo contém tantas subdivisões quantas as fases intermediárias. Essas subdivisões podem ser de ordem fónica, morfológica ou semântica. Vamos ver como se comporta o português dentro destas subdivisões.

3.2. Os casos de mudança fonética/fonológica

Dos 64 artigos com representação do português, 9 comportam subdivisões de ordem fónica⁵. Observemos o caso do artigo */ro'tund-u/ (Hegner, in DÉRom). Divide-se em 3 grandes subdivisões, que correspondem a três tipos, o primeiro deles dividindo-se, por suas vez, em dois subtipos: I. 1. Tipo arcaico original */ro'tund-u/ ; I. 2. Tipo arcaico metatético */to'rund-u/; II. Tipo aferético */'tund-u/ ; III. Tipo dissimilado */re'tund-u/. O português *redondo* provém do tipo dissimilado */re'tund-u/, o mais difundido e também o mais recente, como se vê pela sua ausência em sardo : o étimo */re'tund-u/ é pois posterior à 2ª metade do séc. II (data da separação do protossardo do ramo românico, segundo Straka, 1956: 256). Ora, como étimo de *redondo*, o DELP₃ dá “lat. *rotundu-*”; o Houaiss, porém, propõe uma forma hipotética **retundus* que coincide com a nossa, porém com a marca da declinação latina (neste aspeto, o DELP₃ é mais rigoroso). No caso de port. *neve*, o DELP₃ dá como étimo “Do lat. *nīve-*” e o Houaiss: “lat. *nix, nivis* ‘neve’”. Como se explicaria então a vogal aberta do português? DENF é o único dos três que menciona um étimo mais coerente com o lexema português: “Do lat. vulg. *něvem* (cláss. *nix, nīvis*)”. No artigo */nīβ-e/ (Delorme, in DÉRom), consideram-se duas subdivisões fónicas para este lexema

⁵ */βad-u/; */eder-a/; */erb-a/ ~ */erβ-a/; */ϕak-e-/; */ϕen-u/ ~ */ϕen-u/; */kas'tani-a/ ~ */kas'tini-a/; */kuer-e-/ ~ */kuer-e-/; */nīβ-e/; */rotund-u/.

panromânico: I. Tipo original */niβ-e/; II. Tipo apresentando uma atração paronímica */neβ-e/. O tipo II., cuja abertura da vocal tónica se explica pela atração paronímica de */neβul-a/⁶, cobre uma zona extensa e contínua desde o sudoeste da Romania (occit., gasc., esp., ast., gal./port.), ao passo que o tipo I está representado em todos os outros idiomas (roum., dálm., istriot., it., sard., friul., lad., romanch., fr., frpr.), e os dois tipos sobrepõem-se na maior parte da Itália e em dalmata. A ausência do tipo II. em sardo e em romeno permite datar o surgimento deste tipo do séc. IV, já que é posterior à separação do protorromeno da Dácia (fim do séc. III, cf. Straka, 1956: 258).

3.3. Os casos de mudança morfológica

13 dos 64 artigos do DÉRom comportam subdivisões morfológicas. Em sete casos, as subdivisões refletem uma mudança de género⁷. O artigo */sal-e/ (Yakubovich in DÉRom), por exemplo, comporta três subdivisões, que correspondem às três fases pelas quais se efetuou esta mudança de género entre o latim e as línguas românicas: I. “Substantivo masculino original”; II. Substantivo feminino inovador”; III. “Substantivo masculino restaurado”. O português está representado na terceira subdivisão. Com efeito, o género masculino do lexema português não vem diretamente do masculino latino. A reconstrução protorromance mostrou que o sardo é o único que conservou este género original e que, mais tarde, a maioria dos idiomas românicos (dialetos romenos, dialetos italianos setentrionais, frpr., occit., gasc., cat., esp. e ast.) inovaram passando para o feminino. Só numa terceira fase é que o masculino inicial foi restaurado em alguns idiomas (istriot., friul., lad., romanch., fr., ast., gal. e port.), conservando o asturiano os dois géneros. O lexema */sal-e/ faz parte de um grupo de substantivos originalmente masculinos, que passaram para o feminino já em época protorromance numa grande parte da Romania, seguindo a tendência analógica do latim falado a feminizar os substantivos da terceira declinação (cf. Dardel, 1976). Assim, temos port. *leite*, *mel*, *sangue*, *fel*, *mar*, todos masculinos (restaurados), mas esp., *leche*, *miel*, *sangre* femininos e *fiel* e *mar* que possuem os dois géneros. Sem a comparação românica e a reconstrução, seria difícil conhecer a história da mudança de género deste lexema em português. Por isso é que os dicionários

⁶ Para as hipóteses que imputam a vogal aberta à atração de */gɛl-u/ ou à analogia com */lɛβ-e/, cf. Delorme, in DÉRom, s.v. */niβ-e/, n. 1.

⁷ */ali-u/ ; */βad-u/ ; */lakt-e/ ; */mɔnt-e/ ; */pan-e/ ; */pɔnt-e/ ; */sal-e/.

etimológicos não mencionam esta evolução (DELP₃) ou dizem apenas que o género é masculino como em latim (Houaiss)⁸.

Outro caso de mudança morfológica é o da mudança de classe flexional que afeta os verbos, por exemplo o verbo *fugir*. O artigo */ϕug-e-/ (lat. cl. *fugĕre*) comporta duas subdivisões morfológicas: I. “Flexão em */-e-/” e II. “Flexão em /-i-/”. O continuador português *fugir* (tal como cat., esp., ast. e gal.) encontra-se na segunda subdivisão. A reconstrução mostra que o tipo em /-i-/ , */ϕug-i-re/, donde vem o verbo português, já existia em época protorromance. Revela um primeiro tipo morfológico, minoritário (it. setentrional e central, dialetos sardos arcaicos, gasc., cat. norte-oriental) – sempre em concorrência com um representante do segundo tipo –, e um segundo tipo, que abrange o conjunto do espaço coberto pelo lexema */ϕug-/. A repartição espacial mostra que */ϕug-e-re/ é recessivo e, portanto, constitui a camada mais antiga, enquanto */ϕug-i-re/, geograficamente extensivo, provém de uma inovação protorromance. Nas fontes escritas, *fugere* está documentado desde Plauto [* 254 - † 184], ao passo que *fugire* não aparece nos textos antes de 343/350 (cf. (Jatteau in DÉRom, s.v. */ϕug-e-/). Esta mudança de classe flexional, que observamos no verbo português *fugir* é, porém, bastante antiga se a compararmos com aquela que apresenta o verbo *cair*. Com efeito, apesar de o étimo protorromance */kad-e-/ (lat. cl. *cadĕre*) ter conhecido duas formas, */kad-e-re/, antigo e recessivo, e */ka'd-e-re/, inovador e extensivo, que deu o português antigo *caer*, a forma em *-ir* não remonta além da época medieval (cf. Buchi, in DÉRom s.v. */ka'd-e-/; Malkiel, 1986; Piel, 1989, p. 216; Williams, 1991, p. 166 [§ 148]). Assim, se ambos os verbos *fugir* e *cair* sofreram uma mudança de classe flexional, esta mudança não se efetuou na mesma época para os dois verbos.

Para concluirmos os casos de mudança morfológica, vejamos como a reconstrução românica permite explicar as formas peculiares que revestem o futuro e o condicional do verbo *fazer* e inscrevê-las num processo comum a várias línguas românicas. O artigo */ϕak-e-/ (Buchi, in DÉRom; lat. cl. *facĕre*) evidencia dois tipos morfológicos protorromances, refletidos pelas duas subdivisões do artigo: I. */ϕak-e-re/, original; II.

⁸ Ver também o caso de port. *ponte*. O lexema protorromance conheceu três estados sucessivos : I. masc. isolado no sardo ; II. fem. em áreas laterais e isoladas, entre as quais o port.; III. masculino inovador na Romania central (cf. Andronache in DÉRom, s.v. *ponte*). O que traz uma resposta ao DELP₃ que diz : “[...] não se apresentou ainda uma hipótese convincente que procurasse explicar o motivo da mudança de género deste voc. em alguns romances, entre os quais o Port. O facto é que no séc IX já se dizia *illa ponte* [...]”.

*/ ϕ a-re/, evoluído. O tipo I., que pertence ao estrato mais antigo do protorromance, cobre praticamente o conjunto da Romania, com exceção de istriota, dalmata, friulano, ladino e romanche. O tipo II. pertence a um estrato mais recente: não teve continuadores em sardo nem em romeno, portanto é posterior à separação do protorromeno, que se efetuou na 2^a metade ou no fim do séc. III. (cf. Straka, 1956: 256; Rosetti, 1986: 184). Está representado na Romania central e ocidental – e portanto em português –, onde os dois tipos apresentam, seguindo diferentes esquemas, uma distribuição complementar no seio dos seus paradigmas flexionais, o que testemunha a sua monogénese⁹. O tipo II. procedeu do tipo I. por síncope da sílaba /-ke-/, notadamente na posição proclítica do tipo mais antigo do futuro românico */ ϕ -a-re-'a β -e-/ (cuja extensão é quase totalmente homotópica com a área onde está representado o tipo II.), a frequência do verbo justificando o seu desgaste. Explica-se pois desde a origem como uma variante combinatória do tipo I., que pode também ter sofrido a influência analógica dos verbos */d-a-/ (> port. *dar*) e */st-a-/ (> port. *estar*).

3.4. A variação semântica

Se a maior parte dos étimos reconstruídos no DÉRom são monossémicos, em alguns casos a comparação entre os cognatos românicos incita a pensar que o étimo protorromance conhecia vários sentidos. A reconstrução semântica dos étimos polissémicos, etapa mais delicada na redação dos artigos (cf. Buchi, 2012), leva outra vez a subdivisões. 14 dos 64 artigos apresentam subdivisões que traduzem uma variação semântica¹⁰. É preciso aqui dizer que só interessam para o nosso dicionário as mudanças de sentido dos cognatos que são úteis à reconstrução do sentido do lexema protorromance ou de sub-protorromances regionais, ou seja, as que são mais ou menos comuns aos idiomas românicos ou a subgrupos de idiomas românicos. Não se tomam em conta os sentidos secundários que se desenvolveram num ou noutro idioma românico (cf. LB: 47).

Tomemos o português *barba*. O artigo *barba* no DELP₃ apresenta-se assim: “**Barba**, s. Do lat. *barba* “barba (do homem e dos animais); *fig.*, ramos novos, folhas tenras, penugem” [...].” O Houaiss retoma, na parte da etimologia, essas palavras do DELP₃: “lat. *barba,ae* ‘barba (do homem ou dos animais)’[...]. Neste artigo como nos outros do

⁹ Para a explicação monogenética proposta por Ascoli e Meyer-Lübke e mais ou menos aceite por Corominas, que dá o tipo latino *fagere* como origem das formas em *far*, ver a refutação desenvolvida em Buchi, in DÉRom, s.v. */ ϕ ak-e-/, n. 16. Para mais explicações sobre o desenvolvimento do tipo II., ver a parte “Comentário” no mesmo artigo.

¹⁰ */anim-a-; */as'kult-a-; */barb-a'/; */ β indik-a-; */es'kult-a-; */ ϕ amen/; */klam-a-; */kresk-e-; */kuer-e- ~ */kuer-e-; */le β -a-; */ment-e-; */sa'gitt-a-; */ti'tion-e-; */unkt-u/.

DELP₃, não se dá o sentido do lexema português, o que implica tacitamente que este deveria ter o mesmo sentido que o do étimo latino (clássico) sugerido. Já o DENF é mais preciso: “**barba**, *sf.* ‘cabelos do rosto do homem’ XIII. Do lat. *barba-ae* [...]”. Mas desta vez, é o latim que não tem definição. No primeiro caso, a pergunta é: qual é o sentido do port. *barba*? No segundo: qual é o sentido do latim *barba*? No português contemporâneo, este lexema é polissêmico: o DLP conta 6 aceções, o DLPC 8 e o Houaiss 14. Para qual ou quais delas remete o DELP₃? E o latim só teria a aceção de ‘cabelo do rosto do homem’? O sentido “queixo” não está documentado no latim escrito da Antiguidade, o que não significa que não existisse já no latim falado. A reconstrução românica leva a um étimo protorromance bissêmico e o artigo do DÉRom distingue os dois sentidos com duas subdivisões: I. “Barba”; II. “Queixo”. Se o primeiro é panromânico, o segundo está também representado de leste a oeste da Romania (dacorom., istrorom., dálm., istriot., it., sard., friul., occit., cat., esp., ast., gal. e port.), e isso desde a primeira fase do protorromance, antes da separação do sardo (2^a metade do séc. II, *cf. supra*). O REW₃, que também dá apenas o sentido “*Bart*” [“barba”] para o latim (como aliás o FEW [1: 243b]), já assinala: “*Die Bedeutung ‚Kinn‘ [“queixo”] ist über fast ganz Italie und die Iberische Halbinsel verbreitet*”. Ora, onde aparece o sentido “queixo” nos dicionários etimológicos do português? Nem em português, nem em latim. Apenas encontramos rasgos desta aceção no DENF, no derivado *abarbar*, documentado no séc. XIII, e dado com o significado de “tocar com a barba ou queixo”. Para evitar toda a ambiguidade na interpretação dos significados, os lemas do DÉRom não contêm apenas uma definição sintética do lexema protorromance e sim uma definição componencial, muito precisa, que não dá lugar a dúvidas. Assim, os significados sintéticos “barba” e “queixo” encontram-se explicitados na definição proposta no lema: “*ensemble des poils qui poussent au bas du visage de l’homme (sur le menton et les joues); partie du visage située sous la lèvre inférieure et constituée par l’extrémité du maxillaire inférieur*” (Schmidt, in DÉRom, s.v. */barb-a¹).

O caso de *logo* também é interessante. O português antigo conhecia este advérbio também no seu sentido original de substantivo. A comparação românica mostra que já no protorromance ocidental se desenvolvera um advérbio temporal */lɔk-o/, procedente por conversão do substantivo */lɔk-u/ (fixado no caso oblíquo), que significava num primeiro tempo “no mesmo sítio, aqui” e daí “no momento” e depois “dentro de pouco tempo”, e cujos representantes se confundiam com os de */lɔk-u/ (*cf. Gouvert, in DÉRom, s.v.*

*/lɔk-u/. O surgimento do advérbio provocou, numa vasta área do sudoeste da România, a perda do semantismo do substantivo, cujos descendentes sofreram a concorrência do derivado */lo'k-al-e/ (aoccit., gasc., cat.) ou foram substituídos por ele (esp., gal./port.).

3.5. A datação

O trabalho de revisão dos artigos do DÉRom permitiu-nos comparar, sintetizar, emendar e melhorar as datações propostas no DELP₃, no Houaiss e nos trabalhos de Cunha. Antes de darmos os nossos resultados, apresentemos as regras que se aplicam na datação dos cognatos do DÉRom: “[A] datação só toma em conta as atestações diretas, com exclusão de todo tipo de atestação indireta: as unidades românicas não são datadas através de um derivado ou de um composto, nem através de um nome próprio, nem através de um empréstimo feito por outra língua, nem através de uma atestação de aparência românica recolhida num texto alógloto (notadamente latino ou eslavo)” (LB: 50). No caso do galego e do português, todo o cognato anterior a meados do séc. XIV comporta uma só datação para os dois idiomas. Esta regra é discutível, pelo menos para os textos não literários, mas distinguir a datação dos dois idiomas trazia problemas ainda maiores. A decisão de não considerar as atestações em textos latinos foi tomada porque muitas vezes se torna difícil saber se a palavra é latina ou de outra língua. Perdem-se assim muitas atestações, porque numerosas palavras portuguesas surgem em textos latinos ou “latino-portugueses”, mas ganha-se em fiabilidade, já que “A língua notarial hispânica em geral, e a latino-portuguesa, em particular, documenta uma cadeia ininterrupta de romanceamento gráfico que vai desde os documentos mais antigos conhecidos até aos primeiros textos verdadeiramente românicos do início do século XIII” (Emiliano, 2003: 80).

O DELP₃ dá numerosíssimas primeiras datações em textos latinos: representam 26 dos nossos 64 lexemas. O problema é que raríssimas vezes (apenas em dois casos) propõe também uma primeira datação em textos portugueses. O Houaiss, pelo contrário, se retoma as datações latinas do DELP₃, acrescenta sempre uma data da primeira documentação num texto português. No entanto, o Houaiss não aproveitou todos os materiais do DELP₃: levantámos quatro casos (*ano, parte, queijo, salva*) em que a data que propõe é posterior, sem justificação, à do DELP₃, e oito em que a data é menos exata (*crecer, dez, filho, leite, março, nabo, seta, tição*). Em dois casos (*fava e levar*), o Houaiss utiliza o DELP₃ como

fonte de datação mas engana-se ao retomar a data¹¹. O Houaiss propõe 11 retrodatações em relação ao DELP₃ (sobre os 64 lexemas da nossa amostra). As fontes exclusivas destas retrodatações são, para os lexemas que estudamos, os trabalhos de Cunha: 6 de Cunha 1986-1994, siglado no Houaiss “IVPM”, e 5 do futuro Cunha 2006-2007, a que os autores do dicionário tiveram acesso antes da sua publicação, através do ficheiro manuscrito conservado na Fundação Casa de Rui Barbosa, e que figura no Houaiss com a sigla “FichIVPM”.

Vejam agora as datações e as fontes do DÉRom. Neste dicionário, todas as atestações levam as indicações precisas e verificáveis das fontes e a totalidade dos 64 lexemas portugueses estão datados. Em 9 casos a primeira datação (em textos portugueses) é dada pelo DELP₃ (*chantagem, dez, exir, filho, mão, nabo, queijo, seta, vingar*), em 3 outros pelo Houaiss (*hera, louro, maio*), em 3 mais por Cunha 1986-1994 (*alho, beber, cadeia*) e por Cunha 2006-2007 em 2 casos (*salva, unto*). Notamos que Cunha 2006-2007 não recolhe todos os dados de Cunha 1986-1994: as palavras *alho, cadeia, crescer* levam em Cunha 2006-2007 as respetivas datas de 1269, séc. 13 e séc. 13, e em Cunha 1986-1994, 1254, 1265 e 1264. Duas datações foram recolhidas de *Leges* (*mosto, pão, parte*). Vemos que o DELP₃ continua a ser uma fonte indispensável para a datação do léxico português. 20 dos nossos 64 lexemas têm como documento de primeira atestação uma das fontes que acabamos de citar. E os outros 44? Todos esses 44 lexemas têm como fonte de primeira documentação o material recente publicado na Galiza em versão informatizada: o DDGM (*Diccionario de Dicionarios do Galego Medieval*) e o TMILG (*Tesouro Medieval Informatizado da Lingua Galega*). 23 dos 44 lexemas encontram-se no primeiro, 21 no segundo. Podemos assim avaliar a importância dessas fontes galegas que nos permitiram retrodatar 68,75% dos nossos lexemas. É claro que, como sempre em etimologia, essas datações são provisórias.

O DÉRom também dá, quando possível, a datação de semantismos, como no caso dos verbos *ascoitar, chamar, escutar e levar*. E ainda, sempre que possível, a data de última documentação de um lexema ou de um semantismo. Por exemplo, *leixar*, que cedeu o passo a *deixar*, não se encontra – nas fontes bibliográficas consultadas que sempre podem ser aumentadas – depois de 1552; *chantagem* (planta plantaginácea,) desapareceu dos textos após o séc. 16 e foi substituída pela forma metatética *tanchagem*; *levar-se* no sentido de “levantar-se” não se encontra depois do séc. 14.

¹¹ Dá as datas de 1114 para *fava* e 1022-1055 para *levar*, em vez de, respetivamente, 1141 e 1055-1065.

3.6. A correção de erros da lexicografia

Além das numerosas correções de datação que acabamos de ver, e que não são propriamente erros senão falta de atualização normal em obras que não dispunham dos meios tecnológicos que temos hoje, a confrontação das diversas fontes consultadas para o português permitiu corrigir erros que aparecem na lexicografia galega, portuguesa e brasileira. Começando pelos mais leves, pudemos emendar referências erradas, como na paginação dada pelo DELP₃ quando se refere aos *Opúsculos* de Vasconcelos, para o lexema *feno* (cf. Benarroch, no prelo 1). Depois corrigiram-se erros vindos de uma leitura um pouco rápida dos textos utilizados como fontes. Por exemplo, erros tipográficos com consequências fónicas: a forma *rodonho*, mencionada no DELP₃, s.v. *redondo*, como documentada em *Diplomata et Chartae*, 1, figura neste documento sob a forma *rodondo*. Ou ainda erros de avaliação do sistema linguística do texto, provocando erros de datação: assim, o texto de 1192 (*O Auto de partilhas*) citado por DELP₃ e DDGM como primeira documentação de *parte* não é senão uma tradução realizada provavelmente no fim do séc. XIII de um texto latino de 1192 (cf. Costa, 1992: 167-256; Castro, 2004: 75). E, para acabar, erros de homonímia, por exemplo, as duas atestações dadas por Cunha 2006-2007 para *salva* (s.f. «planta») não se referem à planta */salβi-a/ (*salvia*), mas são formas femininas do adj. *salvu*.

4. Conclusão

A reconstrução românica permitiu precisar as filiações que levaram do latim ao português, revelando étimos latinos com significantes mais coerentes com as formas fónicas do português, esclarecendo casos de mudança morfológica (mudança de género e de classe flexional) e pondo em relevo semantismos do latim falado e evoluções semânticas. Estes aspetos do latim falado, refletidos pelo protorromance reconstruído, nem sempre estão documentados no latim escrito da Antiguidade e não raramente faltam nos dicionários etimológicos portugueses, quando são essenciais para a *etimologia proxima*. Graças à utilização de fontes galegas até então não aproveitadas pela etimologia portuguesa, foi possível retrodatar perto de 70% dos 64 lexemas portugueses estudados. A recolha de informações até então espalhadas em diversas fontes assim como o confronto entre estas fontes permitiram sintetizar os conhecimentos existentes sobre esses lexemas e corrigir certos erros. Com estes resultados, esperamos ter demonstrado os aspetos inovadores do DÉRom e a sua contribuição para uma etimologia da língua portuguesa, de qualidade, viva e atualizada.

Referências

- Andronache, M. (no prelo) Le statut des langues romanes standardisées contemporaines dans le DÉRom. In Casanova Herrero, Emili e Calvo Rigual, Cesáreo (eds.), *Actes del 26^e Congrès Internacional de Lingüística i Filologia Romàniques (València, 6-11 de setembre de 2010)*. Berlim: De Gruyter, 4.
- Baldinger, K. (1959) L'étymologie hier et aujourd'hui. *Cahiers de l'AIEF* 11, pp. 233-264.
- Benarroch, M. (2010) L'apport des dictionnaires de Jerónimo Cardoso (XVI^e siècle) à la datation du Dicionário Houaiss (2001). In Iliescu, Maria, Siller-Runggaldier, Heidi e Danler, Paul (éd.), *Actes du XXVe Congrès International de Linguistique et de Philologie Romanes (Innsbruck, 3-8 septembre 2007)*. Berlim e Nova Iorque: de Gruyter, 2, pp. 623-632.
- Benarroch, M. (no prelo 1) L'apport du DÉRom à l'étymologie portugaise. In Casanova Herrero, Emili e Calvo Rigual, Cesáreo (eds.), *Actes del 26^e Congrès Internacional de Lingüística i Filologia Romàniques (València, 6-11 de setembre de 2010)*. Berlim: De Gruyter, 4.
- Benarroch, M. (no prelo 2) L'étymologie du lexique héréditaire : en quoi l'étymologie panromane est-elle plus puissante que l'étymologie idioromane ? L'exemple du DÉRom (*Dictionnaire Étymologique Roman*). Chabrolle-Cerretini, Anne-Marie (éd.), *Actes du colloque "Romania : réalité(s) et concepts", Nancy, 6-7 octobre 2011*. Limoges: Lambert et Lucas.
- Benarroch, M. (no prelo 3) Latin oral et latin écrit en étymologie romane : l'exemple du DÉRom (*Dictionnaire Étymologique Roman*). In Araújo Carreira, Maria Helena (ed.), *Actes du colloque "Les rapports entre l'écrit et l'oral dans les langues romanes" (Université Paris 8 Vincennes-Saint-Denis, 9/10 décembre 2011)*.
- Boullón Agrelo, A. I. (ed.) (2007) *Na nosa lyngoage galega. A emerxencia do galego como lingua escrita na Idade Media*. Santiago de Compostela: CCG e ILG.
- Buchi, É. (2010) Where Caesar's Latin does not belong: a comparative grammar based approach to Romance etymology. In Brewer, Charlotte (ed.) *Selected Proceedings of the Fifth International Conference on Historical Lexicography and Lexicology held at St Anne's College, Oxford, 16-18 June 2010*, Oxford, Oxford University Research Archive (<http://ora.ox.ac.uk/objects/uuid%3A237856e6-a327-448b-898c-cb1860766e59>).

- Buchi, É. (2012) Des bienfaits de l'application de la méthode comparative à la matière romane : l'exemple de la reconstruction sémantique. In Vykypěl, Bohumil e Boček, Vít (ed.) *Methods of Etymological Practice*. Praga: Nakladatelství Lidové noviny, pp. 105-117.
- Buchi, É. e Schweickard, W. (2009) Romanistique et étymologie du fonds lexical héréditaire : du REW au DÉRom (*Dictionnaire Étymologique Roman*). In Alén Garabato, Carmen et al. (ed.), *La Romanistique dans tous ses états*. Paris: L'Harmattan, pp. 97-110.
- Buchi, É. ; Schweickard, W (2010) À la recherche du protoroman: objectifs et méthodes du futur Dictionnaire Étymologique Roman (DÉRom). In Iliescu, Maria, Siller-Runggaldier, Heidi e Danler, Paul (éd.), *Actes du XXV^e Congrès International de Linguistique et de Philologie Romanes (Innsbruck, 3-8 septembre 2007)*. Berlin e Nova Iorque: de Gruyter, 6, pp. 61-68.
- Campbell, L. (2004² [1998¹]) *Historical Linguistics. An Introduction*. Cambridge: MIT Press.
- Castro, I. (2004) A primitiva produção escrita em português. In *Orígenes de las lenguas romances en el Reino de León. Siglos IX-XII*. León: Centro de Estudios e Investigación San Isidoro, pp. 69-97.
- Castro, I. (2008 [¹2004]) *Introdução à história do português. Geografia da língua. Português vivo*. Lisboa: Colibri.
- Chambon, Jean-Pierre (2007) Remarques sur la grammaire comparée-reconstruction en linguistique romane (situation, perspectives). *Mémoires de la Société de linguistique de Paris* 15, pp. 57-72.
- Chambon, Jean-Pierre (2010) Pratique étymologique en domaine (gallo)roman et grammaire comparée-reconstruction. À propos du traitement des mots héréditaires dans le *TLF* et le *FEW*. In Choi-Jonin, Injoo, Duval, Marc e Soutet, Olivier (ed.) *Typologie et comparatisme. Hommages offerts à Alain Lemaréchal*. Lovaina, Paris : Walpole, Peeters, pp. 61-75.
- Chambon, Jean-Pierre (no prelo) Réflexions sur la reconstruction comparative en étymologie romane (entre Meillet et Herman). In Glessgen, Martin-D. e Schweickard, Wolfgang (eds.), *Étymologie romane. Objets, méthodes et perspectives. Actes du Colloque international en l'honneur du Professeur Max Pfister à l'occasion de son 80^e anniversaire, Université de Zurich, 19-21 avril 2012*.

- Costa, P. A. de J. da (1992) *Estudos de cronologia, diplomática, paleográfica e histórico-linguísticos*. Porto: Sociedade Portuguesa de Estudos Medievais.
- Cunha, A. G. da (1986-1994) *Índice do Vocabulário do Português Medieval*. Vol. 1 [A] 1986; vol. 2 [B-C] 1988; vol. 3 [D] 1994. Rio de Janeiro: Fundação Casa de Rui Barbosa.
- Cunha, A. G. da (1992) Aditamento ao Índice do Vocabulário do português medieval. *Confluência* 3, pp. 23-35.
- Cunha, A. G. da (dir.) (2006/2007² [2002¹]) *Vocabulário Histórico-Cronológico do Português Medieval* (DVD). Rio de Janeiro: Fundação Casa de Rui Barbosa e Ministério da Cultura.
- Dardel, R. de (1976) Une analyse spatio-temporelle du roman commun reconstruit (à propos du genre). In Vârvaro, Alberto (ed.), *XIV Congresso Internazionale di Linguistica e Filologia Romanza, Napoli 15-20 aprile 1974*. Napoli, Amsterdam: Macchiaroli, Benjamins, 14/2, pp. 75-82.
- Dardel, R. de (2007) Une mise au point et une critique relatives au protoroman. *Revue de linguistique romane* 71, pp. 329-357.
- Dardel, Robert de (2009) La valeur ajoutée du latin global. *Revue de linguistique romane* 73, pp. 5-26.
- DCECH = Corominas, Joan e Pascual, José Antonio (1980-1991) *Diccionario crítico etimológico castellano e hispánico*, 6 vols. Madrid: Gredos.
- DDGM = González Seoane, Ernesto, Álvarez de la Granja, María e Boullón Agrelo, Ana Isabel (2006) *Diccionario de diccionarios do galego medieval*. Santiago de Compostela: Universidade de Santiago de Compostela (<http://sli.uvigo.es/DDGM/>).
- DELP₃ = Machado, José Pedro (1977³ [1952¹]) *Dicionário etimológico da língua portuguesa*, 5 vols. Lisboa: Horizonte.
- DENF = Cunha, Antônio Geraldo da (1982) *Dicionário Etimológico Nova Fronteira da língua portuguesa*. Rio de Janeiro: Nova Fronteira.
- DÉRom = Buchi, Éva e Schweickard, Wolfgang (dir.) (2008-) *Dictionnaire Étymologique Roman* (DÉRom). Nancy: ATILF (<http://www.atilf.fr/DERom>).
- DLP = *Dicionário da Língua Portuguesa* (2005) (DVD). Porto: Porto Editora
- DLPC = *Dicionário da Língua Portuguesa Contemporânea* (2001), 2 vols. Lisboa: Academia das Ciências de Lisboa e Verbo.
- Emiliano, A. (2003a): Sobre a questão d'‘os mais antigos textos escritos em português’’. In Castro, Ivo e Duarte, Inês (eds.), *Razões e Emoção. Miscelânea de estudos em*

- homenagem a Maria Helena Mira Mateus*. Lisboa: Imprensa Nacional-Casa da Moeda, pp. 261-278.
- Emiliano, A. (2003b) Os estudos dos documentos notariais latino-portugueses e a história da língua portuguesa. *Signo. Revista de Historia de la Cultura Escrita*, 11, pp. 75-122.
- FEW = Wartburg, Walther von *et al.* (1922–2002) *Französisches Etymologisches Wörterbuch. Eine Darstellung des galloromanischen Sprachschatzes*, 25 vols. Bonn, Heidelberg, Berlin, Basileia: Klopp, Winter, Teubner, Zbinden.
- Fischer, I. (1969) III. Lexicul. 1. Fondul panromanic. In Rosetti, Alexandru *et al.*, *Istoria limbii române*. Bucureste: Editura Academiei Republicii Socialiste România, 2, pp. 110-116.
- Houaiss, A.; Villar, M. de S. (2001) *Dicionário eletrônico Houaiss da língua portuguesa* (DVD, versão 2.0a). Rio de Janeiro: Objetiva.
- LB = *Livre Bleu* (versão em linha do 03/01/2013). *DÉRom* (Dictionnaire Étymologique Roman). Nancy: ATILF, <http://www.atilf.fr/DÉRom>.
- Leges* = Herculano, Alexandre (dir.) (1856/1868). *Portugaliae Monumenta Historica a Saeculo Octavo post Christum usque ad Quintum Decimum. Leges et Consuetudines*, 2 vols. Lisboa : Academia Real das Ciências de Lisboa.
- LEI = Pfister, Max e Schweickard, Wolfgang (dir.) (1979–) : *Lessico Etimologico Italiano*. Wiesbaden: Reichert.
- Lisboa, E. de (1937) *O dicionário do Sr. Nascentes e o REW. Rectificações*. Rio de Janeiro: Pimenta de Mello.
- Lorenzo, R. (1968) *Sobre cronologia do vocabulário galego-português (Anotações ao 'Dicionário etimológico' de José Pedro Machado*. Vigo: Galaxia.
- Maia, C. de A. (1986) *História do Galego-Português. Estado lingüístico da Galiza e do Noroeste de Portugal desde o século XIII ao século XVI (com referência à situação do galego moderno)*. Lisboa: Fundação Calouste Gulbenkian e Junta Nacional de Investigação Científica e Tecnológica.
- Malkiel, Y. (1986) From Old Portuguese *caer* “to fall” to Modern *cair* : Three Explanations. *Romance Quarterly* 33, pp. 5-10.
- Martins, A. M. (1999) Os mais antigos textos escritos em português: Documentos de 1175 a 1252. In Hub Faria, Isabel (ed.), *Lindley Cintra. Homenagem ao Homem, ao Mestre e ao Cidadão*. Lisboa: Cosmos e Faculdade de Letras da Universidade de Lisboa, pp. 491-534.

- Martins, A. M. (2007) O primeiro século do português escrito. In Boullón Agrelo, Ana Isabel (ed.), *Na nosa lyngoage galega. A emerxencia do galego como lingua escrita na Idade Media*. Santiago de Compostela: CCG / ILG, pp. 161-184.
- Mattos e Silva, R. V. (2008) *O português arcaico: uma aproximação*, 2 vols. I: Léxico e morfologia; II: Sintaxe e fonologia. Lisboa: Imprensa Nacional-Casa da Moeda.
- Meier, H. (1964) Zur Geschichte der romanischen Eymologie. *Archiv für das Studium der neueren Sprachen und Literaturen* 201, pp. 81-109.
- Monjour, A. (2004) El diccionario Houaiss y la etimología portuguesa. In Boullón Agrelo, Ana Isabel (ed.), *Novi te ex nomine. Estudos filolóxicos ofrecidos ao Prof. Dieter Kremer*. A Coruña: Fundación Pedro Barrié de la Maza, pp. 145-155.
- Nascentes, A. (1932) *Dicionário etimológico da língua portuguesa*. Rio de Janeiro: Livraria Francisco Alves *et alii*.
- Pfister, M. e Lupis, A. (2001) *Introduzione all'etimologia romanza*. Soveria: Manelli, Rubbettino.
- Piel, J. M. (1932) Notas à margem do “Romanisches Etymologisches Wörterbuch”. *Biblos* 8, pp. 379-392.
- Piel, Joseph M. (1989) *Estudos de linguística histórica galego-portuguesa*. Lisboa : Imprensa Nacional.
- REW₃ = Meyer-Lübke, Wilhelm (1935³) *Romanisches Etymologisches Wörterbuch*. Heidelberg: Winter.
- Rosetti, A. (1986) *Istoria limbii române. De la origini și pînă la începutul secolului al XVII-lea*. Bucureste: Editura Științifică și Enciclopedică.
- Schweickard, W. (2012) Le Dictionnaire Étymologique Roman (DÉRom) entre tradition et innovation. In Trotter, David, (ed.), *Present and future research in Anglo-Norman : Proceedings of the Aberystwyth Colloquium, 21-22 July 2012*. Aberystwyth : The Anglo-Norman Online Hub, pp. 173-178.
- Souto Cabo, J. A. (2006) Inventário dos mais antigos documentos galego-portugueses. *Agália* 85-86, pp. 9-88.
- Straka, G. (1956) La dislocation linguistique de la Romania et la formation des langues romanes à la lumière de la chronologie relative des changements phonétiques. *Revue de linguistique romane* 20, pp. 249-267.
- TMILG = Varela Barreiro, Xavier (2004–) *Tesouro Medieval Informatizado da Lingua Galega*. Santiago de Compostela: Instituto da Lingua Galega (<http://ilg.usc.es/tmilg/>).
- Vasconcelos, J. L. de (1928) *Opúsculos* I. Coimbra: Imprensa da Universidade.

William, E. (1991⁵ [1961]) *Do latim ao português. Fonologia e morfologia históricas da língua portuguêsã*. Rio de Janeiro: Tempo brasileiro.