

A intervenção do editor no registro das palavras
compostas em textos do português arcaico:
algumas reflexões sobre o uso do hífen

*Editor's intervention in the register of compounds in archaic
Portuguese texts: some reflections on the use of hyphen*

Antonia Vieira dos Santos *
Universidade Federal da Bahia, Salvador, Bahia, Brasil

Resumo: O emprego do hífen no interior das palavras compostas não se desenvolve antes do século XIX. Dessa forma, os compostos que aparecem grafados com esse sinal em textos portugueses arcaicos resultam de critérios de edição adotados pelos editores. A partir desse tipo de intervenção, discute-se a relação entre o uso do hífen – resultado de uma convenção ortográfica – e a definição de composto adotada pelo editor, sendo a sua ausência significativa para a história da língua.

Palavras-chave: Palavras compostas; Português arcaico; Edição de textos; Emprego do hífen.

Abstract: The use of hyphen within compounds was not built up before the nineteenth century. Thus, compounds spelled with hyphen in archaic Portuguese texts are a result of editing criteria adopted by editors. From this type of intervention, the relationship between hyphen use – result of a spelling convention – and the definition of compound adopted by the editor are discussed. Its absence is significant to the history of the language.

* Professora do Departamento de Letras Vernáculas do Instituto de Letras da Universidade Federal da Bahia – UFBA, Salvador, Bahia, Brasil; toniavieira@gmail.com

Keywords: Compounds; Archaic Portuguese; Text editing; Use of hyphen.

1 INTRODUÇÃO

As edições de natureza filológica podem ser caracterizadas pelo grau de intervenção do editor no texto do documento. Assim, há edições que tendem a se manter mais próximas das características originais do texto, no que diz respeito, por exemplo, à ortografia, a fronteiras de palavra, a abreviaturas, à pontuação etc., como as paleográficas ou diplomáticas, ou que trazem maior intervenção do editor, correspondendo a edições críticas, segundo Mattos e Silva (2008, p. 51).

No caso específico da edição de textos do português arcaico, algumas condições têm de ser consideradas para que as edições feitas contribuam, efetivamente, para o conhecimento da língua do período medieval. As interferências processadas pelo editor – em geral explicitadas nos critérios editoriais – terminam por incidir sobre as características linguísticas do texto, não constituindo, para alguns tipos de estudo, como os de natureza fônico-grafemática, por exemplo, edições apropriadas. Portanto, antes de se empreender o estudo de um fenômeno linguístico sobre um determinado aspecto da língua, – fonético-fonológico, morfológico, sintático, semântico e lexical –, faz-se necessário observar os critérios editoriais apresentados pelo editor.

Para um estudo de natureza morfolexical, como é o caso da análise de palavras compostas, a questão é perceber que tipo de intervenção compromete as conclusões derivadas desse estudo. Podem ser referidos, como aspectos morfológicos, os elementos constituintes de uma palavra, como radical (ou semantema), tema, afixos derivacionais, desinências nominais e verbais. No entanto, o elemento que será abordado neste trabalho – o hífen – não integra, de fato, a estrutura morfológica da palavra, uma vez que está circunscrito ao convencionalismo da escrita.

O hífen, segundo o dicionário Houaiss (2009), é um “sinal em forma de um pequeno traço horizontal (-), us. para unir os elementos de palavras compostas, separar sílabas em final de linha e marcar ligações enclíticas e mesoclíticas”, como, por exemplo, em <guarda-chuva> <aboli-/ção> <telefonaram-lhe> <fá-lo -ei>. O mesmo Houaiss estabelece o seu surgimento na língua portuguesa no ano de 1576, especificando a sua etimologia a partir do gr. *huphén*, adv. ‘juntamente’, através do latim tardio *hyphen*.

Antes do novo acordo ortográfico, firmado em 1990 entre os países de língua oficial portuguesa e em vigor desde 2009, o hífen funcionava como indicador do estatuto morfológico de uma palavra composta e como critério para o seu registro como lema em dicionários. As combinações de palavras registradas

modernamente no dicionário sem o hífen não corresponderiam a palavras compostas, mas a locuções, inserindo-se geralmente dentro da estrutura de um verbete. Poder-se-ia, com base nesse critério lexicográfico, estabelecer fronteiras entre compostos e não compostos.

Com o acordo ortográfico, admite-se que o hífen já não constitui parâmetro para demarcar uma palavra composta nem para licenciar uma determinada unidade polilexical como lema. Afinal, não faz sentido pensar que o que era considerado composto antes do acordo ortográfico tenha deixado de sê-lo simplesmente pela ausência, na atualidade, do hífen. Por exemplo, a unidade complexa *pé de moleque* (“doce consistente feito de açúcar ou rapadura com amendoim torrado” grafava-se, antes do acordo, *pé-de-moleque*, com hífen, constituindo, dessa forma, uma entrada de verbete (veja-se Houaiss, 2004, *s.v.* pé-de-moleque). Após a implementação das regras do acordo, *pé de moleque* deixou de constituir lema e passou a integrar uma das locuções com *pé*, dentro do verbete “pé” (veja-se Houaiss, 2009, *s.v.* pé). O mesmo ocorre com *maria vai com as outras* (“pessoa sem personalidade, que se deixa facilmente influenciar ou levar pelos outros”), que perdeu o seu estatuto de lema, tendo sido relegada agora à condição de locução no dicionário (vejam-se Houaiss, 2004, *s.v.* maria-vai-com-as-outras, e Houaiss, 2009, *s.v.* maria). Mesmo (e principalmente) com o acordo ortográfico, pode-se dizer que as regras do uso do hífen nas palavras compostas continuam das mais complexas na ortografia portuguesa.

O objetivo deste trabalho é desenvolver reflexões sobre o uso do hífen em um cenário específico: no registro de palavras compostas em edições de textos do português arcaico. Sabendo-se que o uso do hífen no interior das formas compostas não se desenvolve antes do século XIX, busca-se refletir, a partir da observação das unidades plurilexicais registradas com (e sem) hífen, sobre a noção que os editores de texto têm de uma palavra composta. Que leitura se pode fazer da presença do hífen (ou da sua ausência) em uma edição não conservadora? Para o desenvolvimento dessa discussão, faremos referência ao seu uso também no português contemporâneo.

Nas seções seguintes, abordaremos a questão do surgimento do hífen, discutindo o seu papel na atribuição do estatuto de composto a determinados sintagmas da língua. Trataremos, ainda, da composição de palavras e dos compostos, apresentando o fenômeno composicional segundo uma perspectiva mais abrangente. Na seção correspondente à apresentação e à análise dos dados, buscaremos demonstrar que o seu uso por parte dos editores de texto costuma ser irregular e assistemática no caso de alguns padrões compositivos. Defendemos que esse comportamento reflete, em alguma medida, a concepção de palavra composta do editor e toda a complexidade envolvida na sua definição.

2 O SURGIMENTO DO HÍFEN

A origem do uso do hífen na ortografia do português e a sua evolução no contexto da língua portuguesa são questões tratadas por Rita Marquilhas, no artigo “O acento, o hífen e as consoantes mudas nas Ortografias antigas portuguesas”. Segundo Marquilhas (1987, p. 109), “[p]ara os substantivos compostos não entrava ainda qualquer sinal diacrítico no universo gráfico de João de Barros”. Na verdade, em Portugal, o emprego do hífen no interior das formas compostas não se desenvolve antes do século XIX, quando, com o movimento dos “sônicos”, um dos partidos da disputa ortográfica, em oposição aos etimológicos, intenta-se retratar fonograficamente a língua falada (Marquilhas, 1987, p. 113). Nos gramáticos anteriores ao século XIX, o hífen

aparecia nos manuais no capítulo reservado aos sinais de pontuação, geralmente sob o título de “Outros Sinais”, e funcionava quase sistematicamente em relação a três únicos exemplos: *guarda-porta*, *passa-tempo*, *val-verde*; haviam sido dados por Duarte Nunes de Leão, e sob uma diferente morfologia, a de “guarda **v** porta”, “passa **v** tempo”. (Marquilhas, 1987, p. 113)

Ainda de acordo a autora,

[s]ó em 1739, na *Orthographia* de Madureira Feijó, há o testemunho de o hífen ter sido morfológicamente assimilado à linha de seguimento que, desde os copistas medievais, se traçava no fim da regra (-). (Marquilhas, 1987, p. 113)

Especificamente, Feijó (1861[1739], p. 123) estabelece que:

A *Conjunção*, a que os Gregos chamão *Hyphen*, he hum sinal, que se figura como hum v consoante, com huma risquinha antes, e outra depois direitas, deste modo -v-; e serve este sinal para unirmos duas palavras, que per si são separadas, como se forão huma só na pronunçiação; v.g. *Passa-v-tempo*. *Guarda-v-porta*, etc. Hoje para se evitar o trabalho de estarmos figurando este accento, usamos em seu lugar de huma só risquinha no meio das palavras, que se devem unir: v.g. *Passa-tempo*, *Guarda-porta*, etc.

Com a chegada dos sônicos, portanto, é que o hífen ganha papel mais relevante, tendo o seu uso se estendido para além dos exemplos clássicos. A importância do hífen na marcação de vocábulos compostos, nesse período, está relacionada principalmente com a indicação da pronúncia que deve ser feita, como no caso de *bemaventurado* e *bemfeito*, por exemplo. Barbosa Leão (1875, p. 66, *apud* Marquilhas, 1987, pp. 113-114) justifica, portanto, que “n’estes e analogos casos entendemos que a composição deve ser marcada na orthographia, escrevendo-se *bem-feito*, *bem-aventurado*, etc.”. Nas *Bases para a Unificação da Orthografia*, de 1911, estabelecidas a partir da *Orthografia Nacional* de Gonçalves Viana, afirma-se o seguinte sobre o hífen (p. 44):

Êste sinal prende os vocábulos compostos, quando os seus elementos, conservando a acentuação própria, perdem em parte a sua significação primordial; ex.: *mãe-d’agua*, *porta-bandeira*, *água-forte*, *franco-russo*, *madre-pérola*, etc.

Quanto à sua utilização atual, Cunha e Lindley Cintra, na *Nova Gramática do Português Contemporâneo* (1985, pp. 66-69), elencam as seguintes regras: “a) para ligar os elementos de palavras compostas ou derivadas por fixação (...); b) para unir pronomes átonos e verbos (...); c) para, no fim da linha, separar uma palavra em duas partes”. Esses autores (1985, p. 105) apontam que o uso do hífen, no caso das palavras compostas, resulta de uma convenção ortográfica, não havendo, sob essa perspectiva, diferença entre *chapéu-de-sol* e *pai de família*, por exemplo.

No novo acordo ortográfico, entretanto, o hífen é empregado

nas palavras compostas por justaposição cujos elementos (substantivos, adjetivos, numerais ou verbos) constituem uma nova unidade morfológica e de sentido, mantendo o acento próprio, ainda que o primeiro elemento esteja reduzido: *sócio-gerente*, *arco-íris*, *afro-luso-brasileiro* (*Escrevendo pela nova ortografia*, 2009, p. 42).

No caso das estruturas envolvendo um terceiro elemento, como uma preposição, o novo acordo postula a ausência do hífen, com exceção quando designam espécies botânicas e zoológicas ou quando constituem grafias institucionalizadas, ou seja, consagradas pelo uso (cf. *maria vai com as outras* ‘Pessoa de personalidade fraca que é facilmente influenciável pela opinião de outras’ e *maria-sem-vergonha* ‘Erva da fam. das balsamináceas’ (Houaiss, 2009, *s.v.* *maria* e *maria-sem-vergonha*, respectivamente).

Marquilhas (1987, p. 108) informa que o hífen no interior das formas compostas

nunca foi muito relacionado com os outros dois fenómenos [acentos e consoantes cultas], nunca foi sequer muito interrogado na sua forma e funções próprias, e lá foi aparecendo e desaparecendo conforme os hábitos tipográficos de cada época

o que explica, assim, o seu funcionamento assistemático.

Dessa forma, o hífen, apesar de marcar formalmente uma maior integração semântica entre os elementos, não pode ser tomado como critério de identificação (visual) do composto, pois a sua aplicação, ao se dar unicamente na língua escrita, ocorre por vezes de forma aleatória e assistemática.

Na próxima seção, desenvolveremos algumas ideias sobre a definição do composto a partir de registros sobre o seu uso em gramáticas, históricas e prescritivas, questionando a identificação dessa unidade lexical pelo critério gráfico do hífen.

3 COMPOSIÇÃO DE PALAVRAS E COMPOSTOS: ALGUMAS CONSIDERAÇÕES

A composição de palavras configura-se como um dos processos mais produtivos de formação de palavras do português, ao lado da derivação. A noção de palavra composta, no entanto, envolve uma série de questões não muito bem resolvidas no âmbito da linguística, como a distinção entre compostos e derivados, compostos e unidades fraseológicas e entre compostos e colocações. Vários critérios – fonológico, morfológico, sintático e semântico – são invocados com o intuito de delimitar a palavra composta, mas observa-se que, ainda assim, não constitui uma tarefa fácil definir o que é um composto.

Nesse sentido, pode-se falar em várias definições de composição e palavra composta, como apontam Guevara e Scalise (2009, p. 106)¹, dando-se preferência a falar da sua descrição, ao longo do tempo, através de diferentes olhares e perspectivas.

¹ “Generalizing, the many definitions of compound that one finds in the literature are tightly predetermined by the theoretical choices made by the author(s)” (Guevara; Scalise, 2009, p. 106). Tradução livre: “Generalizando, as muitas definições de composto que se encontram na literatura são fortemente determinadas pelas escolhas teóricas feitas pelo(s) autor(es)”.

Para Nunes (1956, p. 386), a composição é a reunião de

duas ou mais palavras, que tendo tido antes vida própria e independente e possuído cada uma delas sua significação especial, vieram por fim a fundir-se por forma tal que dessa fusão resultou uma única, em geral com um só acento e sempre com uma ideia singular.

Para Said Ali (1964, p. 258), de acordo com o registro na sua *Gramática Histórica*, a palavra composta é

o resultado da evolução e fusão semântica destes elementos, devendo-se notar que muitas das atuais palavras compostas, antes de se fundirem semanticamente para representar uma idéia simples, tiveram um período de existência bastante longo em que não se distinguiam de outros grupos sintáticos. Em outros casos o processo da composição efetua-se desde logo ou em tempo muito breve.

Inserida numa óptica prescritiva da língua, a *Nova Gramática do Português Contemporâneo* (1985, p. 104), de Celso Cunha e Lindley Cintra, define a composição como a “união de dois ou mais radicais”, cujo produto “representa sempre uma idéia única e autônoma, muitas vezes dissociada das noções expressas pelos seus componentes”. Os autores veem o produto composicional, do ponto de vista semântico, como frequentemente opaco. Nesse domínio, distinguem compostos justapostos (*beija-flor, bem-me-quer, madre pérola, passatempo*) e compostos aglutinados (*aguardente, pernalta*), bem como as funções desempenhadas pelos elementos constituintes – determinante ou de determinado.

A organização dos tipos se dá, contudo, segundo a classe gramatical dos elementos constituintes²: NN: *manga-rosa, porco-espinho*; NprepN: *chapéu-de-sol, pai de família*; NA: *aguardente, belas-artes*; AA: *azul-marinho, tragicômico*; NumN: *mil-folhas, trigêmeos*; PronN: *meu-bem, Nosso Senhor*; VN: *beija-flor, passatempo*; VV: *correr-corre, vaivém*; AdvA: *bem-bom, sempre-viva*; AdvV (ou A em função adverbial): *bem-aventurar, maldizer*.

Observa-se, na descrição que os referidos gramáticos fazem sobre a composição, a inclusão de esquemas compositivos como NN, NprepN, NA, AA,

2 Adotamos, neste trabalho, as seguintes notações: NN = Nome + Nome, NprepN (Nome + preposição + Nome), NA = Nome + Adjetivo; AN = Adjetivo + Nome; AA = Adjetivo + Adjetivo; NumN = Numeral + Nome; PronN = Pronome + Nome; VN = Verbo + Nome; VV = Verbo + Verbo; AdvA = Advérbio + Adjetivo; AdvV = Advérbio + Verbo.

PronN, NumN, VN, VV, AdvA, AdvV. Determinadas construções, no entanto, em especial NA e NprepN, apresentam, em sua estrutura, algumas propriedades da combinação sintática. A flexão de número, por exemplo, diferentemente do que é esperado de uma unidade morfológica, em que o plural é marcado à direita, ocorre internamente. Trata-se de compostos sintagmáticos, em que estão presentes não morfemas presos, mas unidades lexicais autônomas. Apenas no caso de formas com as estruturas VN, NN, aparentemente anômalas, parece não haver dificuldade na sua classificação como compostos.

O sumário dessas descrições põe em destaque a tendência de se buscar distinguir, no âmbito da composição, dois tipos de compostos, os “perfeitos”, ou aglutinados, e os “imperfeitos”, ou justapostos. Essa distinção é feita com base na estrutura morfofonológica da palavra, em especial, na morfologia flexional e no acento fonológico. No que se refere à unidade semântica, caracterizadora das palavras compostas, destaque-se a reflexão de Said Ali (1964, p. 258), que aponta para o **caráter gradual** do processo de formação de alguns compostos que se constituem por justaposição³. Antes de se consolidar como um composto, com direito a registro no vocabulário da língua, determinada combinação de palavras experimenta uma fase de transição, indício de que ela está em via de se tornar uma palavra composta. Esse caráter gradual, observado nos compostos, dificulta sobremaneira a análise do estatuto dessas unidades: tratar-se-á de verdadeiras palavras compostas ou de simples grupos sintáticos?

De forma geral, pode-se afirmar que a composição consiste na combinação de dois lexemas, podendo constituir esses lexemas palavras ou radicais.

Na senda do trabalho desenvolvido por Ribeiro (2006) sobre os compostos modernos do português europeu, Santos (2009) defende, em sua tese sobre os compostos sintagmáticos nominais do português arcaico, a hipótese do gradualismo sintático e semântico no âmbito da composição, em especial no âmbito das estruturas sintagmáticas VN, NN, NA, AN e NprepN. A aplicação dos traços de fixação e idiomaticidade aos compostos, mais precisamente aos tipos composicionais, permite a construção de escalas de opacidade sintática e semântica. No aspecto configuracional, tem-se o seguinte *continuum*, que parte do mais opaco para o menos opaco: + opaco [compostos morfológicos > VN > NN > AN > NA > NprepN] – opaco (Ribeiro; Rio-Torto, 2010; Santos, 2009, p. 42). No que diz respeito ao aspecto semântico, admite-se que dentro de cada modelo estão

3 Numa concepção mais abrangente do fenômeno composicional, incluem-se, além dos **compostos sintagmáticos** (ou morfossintáticos), que envolvem duas formas livres (*amor-perfeito, couve-flor*), os **compostos morfológicos**, que envolvem os tipos (i) forma presa + forma presa (*funcionada, biblioteca*) e (ii) forma presa + forma livre (*autoestrada, geoeconômico*).

presentes compostos [+] transparentes e [-] transparentes, não estando atrelados, dessa forma, o conceito de composto e a propriedade de opacidade semântica, unicamente.

Em síntese, ao se assumir como compostos construções cuja estrutura apresentam correlato na sintaxe, como NA e NprepN, uma dificuldade se impõe nos estudos dessas unidades: trata-se da distinção que se deve fazer entre um composto e uma frase (ou combinação livre de palavras), ou ainda, nos estudos morfológicos e lexicais atuais, entre um composto e uma unidade fraseológica e entre um composto e uma colocação.

Nesse aspecto, a lexicografia brasileira utiliza o hífen como elemento demarcador de compostos e outras construções polilexicais. Trata-se de um critério meramente gráfico, pautado, portanto, pela convenção ortográfica. Também a opção pelo uso do hífen nos textos arcaicos sinaliza, de alguma forma, uma tentativa de demarcar a palavra composta, diferenciando-a de outros grupos de palavras. Considerando a inexistência do hífen no interior das palavras compostas nos períodos anteriores ao século XIX (Marquilhas, 1987), que critérios os editores de textos arcaicos utilizaram para identificar como compostos as construções que marcaram com hífen, especialmente aquelas com a estrutura NA, AN e NprepN? Parece-nos que tal escolha está alicerçada numa concepção do editor de palavra composta e do mecanismo de formação de palavras denominado composição.

Neste artigo, a recolha dos dados levou em consideração a proposta seguida por Santos (2009) em sua tese de doutorado, ou seja, a concepção do mecanismo da composição como um processo mais amplo, envolvendo compostos com diferentes níveis de estruturação e pelos quais perpassam diferentes graus de composicionalidade semântica. A identificação das formas compostas levou em conta o caráter referencial ou denotacional por elas apresentado e o atendimento a critérios de natureza morfológica e sintática, como integridade lexical (*saca-rolhas* vs. **saca-muitas-rolhas*) e atomicidade sintática (*fim de semana complicada* vs. *fim da semana complicada*) (Rio-Torto; Ribeiro, 2012, pp. 7-8).

Serão apresentados, a seguir, os dados recolhidos do *corpus* e a análise dos compostos sintagmáticos, relativa ao tipo de registro, ou seja, com a presença e sem a presença do hífen.

4 APRESENTAÇÃO E ANÁLISE DOS DADOS

Resultando o uso do hífen de uma convenção ortográfica, a sua aplicação na língua escrita é realizada, por vezes, de forma aleatória e assistemática. Essa “forma aleatória e assistemática”, verificada na contemporaneidade, pode ser observada também nas edições de textos do português arcaico que optaram por inserir o

hífen, como, por exemplo, as *Cantigas d'Escarnho e de Mal Dizer (CEMD)*, textos poéticos produzidos entre fins do século XIII e meados do século XIV, e editados por Rodrigues Lapa; as *Cantigas de Santa Maria (CSM)*, coletânea medieval de poesias em louvor da Virgem Maria, compostas por Afonso X e reunidas no princípio dos anos sessenta do séc. XIII, editadas por Walter Mettmann; a *Crónica de D. Pedro (CDP)*, cujos manuscritos mais antigos datam do início do séc. XVI ou de fins do séc. XV, editada por Giuliano Macchi; *Vida e feitos d'El Rey Dom João Segundo (VFDJ)* e *Trasladação do Corpo d'El Rey Dom João o Segundo (TCDJ)*, integrantes do *Livro das obras de Garcia de Resende*, aparecido em 1545 em letra de forma, editados por Evelina Verdelho. Ressalte-se que esses textos são referidos por Mattos e Silva (2008, p. 51) como “edições críticas de valor e que podem ser utilizadas para o estudo da língua arcaica”. Estas obras, em conjunto com as edições do *Cancioneiro Geral de Garcia de Resende (GR)*, que reúne poesia produzida em Portugal de meados do século XV até 1516, data de publicação, editado por Aida F. Dias, do *Tratado de Tordesilhas (TT)*, documento régio datado de 1494 editado por Clarinda Maia, e d'A Carta de Pero Vaz de Caminha (CC), editada por Celso Cunha, César Cambraia e Heitor Megale formam o corpus utilizado no presente estudo⁴. Todas essas edições apresentam, em maior ou menor grau, algum tipo de intervenção do editor no texto.

No que diz respeito aos critérios de edição, observa-se, especificamente em relação ao hífen:

CSM – não há referência ao uso do hífen no interior das palavras compostas.

CEMD – não há referência ao uso do hífen no interior das palavras compostas.

CDP – não há referência, na explicitação da “organização externa da edição”, ao hífen.

VFDJ e *TCDJ* – “Introduziram-se hífens e apóstrofes de acordo com as normas ortográficas actuais” (p. 3, da edição eletrônica).

GR – “utilizou-se o hífen e o apóstrofo” (p. XVII).

TT – não há referência ao uso do hífen no interior das palavras compostas.

CC – a edição busca seguir as “Normas para transcrição de documentos manuscritos para a história do Português do Brasil”, cuja norma nº 3 determina que “Não

⁴ A seleção dos textos foi feita a partir do *corpus* utilizado por Santos (2009), representativo do ponto de vista da tipologia textual e das sincronias abrangidas (sécs. XIII-XVI). Essa seleção foi condicionada pela presença de compostos grafados com hífen, com exceção do *Tratado de Tordesilhas* e da *Carta de Pero Vaz de Caminha*, cujas edições apresentam um grau mínimo de intervenção dos respectivos editores.

será estabelecida fronteira de palavras que venham escritas juntas, nem se introduzirá hífen ou apóstrofo onde não houver”.

O tipo de análise que se vai proceder possui um caráter essencialmente descritivo, visando a um cotejo entre compostos com hífen e compostos sem hífen. Esse cotejo se dará entre compostos com a mesma estrutura morfológica, ou seja, serão comparados, por exemplo, compostos NN com e sem hífen, compostos VN com e sem hífen etc., coocorrendo, às vezes, numa mesma obra.

Nas *Cantigas de Santa Maria* ocorrem duas estruturas marcadas com o hífen, as quais foram classificadas como compostas: *ifant-abade* [CSM 164.21] e *madre-donzela* ‘a Santa Maria’ [CSM 153.50], além de *ric-om*’ [CSM 234.17], em que o hífen pode estar pelo apóstrofo, indicando uma elisão, e *ricos-omes* [CSM 235.21]. Por outro lado, o hífen não é utilizado em formas como *alcaide mayor* [CSM 329.47], *altar mayor* [CSM 221.50], *pedras marmores* [CSM 342.2], *agua rosada* [CSM 39.37], *ricas donas* ‘esposa, viúva, filha ou sucessora de um rico-homem’ [CSM 212.13], *meio dia* [CSM 345.59], por exemplo⁵.

Nas *Cantigas de Escárnio e de Maldizer* foi registrado *guarda-cós* ‘espécie de casaco que se usa apertado ao corpo’ [CEMD 209.19], com hífen, enquanto *meio dia* [CEMD 189.9], *monteiro maior* ‘oficial de caça do rei’ [CEMD 324.1] e *ricas donas* [CEMD 238.7] foram grafados sem hífen.

Na *Crônica de D. Pedro* encontram-se grafadas com hífen unidades lexicais como *adeantado-moor* [CDP 16.111], *alcaide-moor* [CDP 9.42], *alferez-moor* [CDP 40.11], *alguazill-moor* [CDP 15.20], *beesteiro-moor* [CDP 20.53], *rricos-homêes* [CDP 17.25]. Observe-se que essas unidades grafadas com hífen apresentam uma particularidade: o segundo elemento, de natureza adjetiva, funciona como uma espécie de “coringa” na formação de palavras compostas. O adjetivo “mor” ‘maior’, nos contextos em que são utilizadas essas formas, corresponde a ‘principal’, ‘superior’, ‘chefe’. Não se grafou o hífen, entretanto, em formas como *meio dia* [CDP 22.32], *mea hidade* [CDP 8.16] (cf. Houaiss, 2009: *meio-dia*, *meia-idade*), *pedras ballaises* ‘rubis’ [CDP 33.63], *guerra guerreada* ‘forma de guerra baseada em emboscadas e ataques de surpresa’ [CDP 32.50].

Na edição de Evelina Verdelho do texto *Vida e Feitos de D. João*, encontram-se grafadas com hífen formas como *alcaide-mor* [VFDJ 6728], *alcaydaria-moor* [VFDJ 8035], *camareira-mor* [VFDJ 5295], *camareiro-mor* [VFDJ 8998],

5 Não vamos apresentar, nessa descrição, exemplos de compostos com a estrutura NprepN, tendo em vista que em nenhum dos casos observados há ocorrência do hífen, não sendo possível opor, dessa forma, estruturas com e sem hífen. No Quadro 1, entretanto, apresentamos alguns exemplos desse tipo de estrutura.

capelão-mor [VFDJ 7015], *capitão-mor* [VFDJ 4060], *contador-mor* [VFDJ 4336], *mordomo-mor* [VFDJ 1704], *porteiro-moor* [VFDJ 4607], *sol-posto* [VFDJ 373] (*sol posto*, em CC), mas *capella mayor* [VFDJ 2259], *agoa doce* [VFDJ 7118], *Mar Ouceano* [VFDJ 970], sem hífen.

Na *Trasladação do Corpo d'El Rey Dom João o Segundo*, há a ocorrência de *conde-prior* [TCDJ 73], *altar-mor* [TCDJ 109] e *capela-moor* [TCDJ 188], grafados com hífen.

No *Cancioneiro Geral de Garcia de Resende*, registram-se com hífen *camareiro-mor*, *comendador-moor* [GR2 197], *Coudel-mor* [GR3 290], *estribeiro-moor* [GR3 157], *guarda-roupa* [GR3 51.2], *porco-pisco* 'pisco-de-peito-ruivo' [GR2 209]. Por outro lado, *mal frances* 'sífilis' [GR3 234.2] (cf. Houaiss, 2009, *s.v. mal-francês*) é registrado sem hífen.

Comparem-se esses registros com os compostos coletados nas edições do *Tratado de Tordesilhas* (TT) e da *Carta de Pero Vaz de Caminha* (CC): *almoçace moor* [TT 2.5], *contador moor* [TT 2v.20], *mordomo moor* [TT 6v.8] e *capitam moor* [CC 1r.1], *fura buchos* [CC 1v.1] e *sol posto* [CC 1v.8], todos sem hífen, ou seja, sem a intervenção dos respectivos editores.

Após essa primeira recolha dos compostos registrados com e sem hífen nas edições caracterizadas como críticas, é importante proceder à análise da incidência (ou inclusão) do hífen nos compostos segundo a sua configuração interna.

Primeiramente, observa-se que as estruturas NA envolvendo a unidade 'mor' são, na maior parte dos casos, marcadas com hífen: *adeantado-moor* [CDP 16.111], *alcaide-moor* [CDP 9.42], *alferez-moor* [CDP 40.11], *alguazill-moor* [CDP 15.20], *beesteiro-moor* [CDP 20.53], *alcaide-mor* [VFDJ 6728], *alcaydaria-moor* [VFDJ 8035], *camareira-mor* [VFDJ 5295], *camareiro-mor* [VFDJ 8998], *capelão-mor* [VFDJ 7015], *capitão-mor* [VFDJ 4060], *contador-mor* [VFDJ 4336], *mordomo-mor* [VFDJ 1704], *porteiro-moor* [VFDJ 4607], *altar-mor* [TCDJ 109] e *capela-moor* [TCDJ 188], *camareiro-mor*, *comendador-moor* [GR2 197], *Coudel-mor* [GR3 290], *estribeiro-moor* [GR3 157]. Poder-se-ia considerar prototípica essa estrutura, N + mor (NA), não oferecendo, para o editor, dificuldades na sua identificação como um nome composto. Por outro lado, não se registra o hífen em *alcaide mayor* [CSM 329.47], *altar mayor* [CSM 221.50], *monteiro maior* [CEMD 324.1], *capella mayor* [VFDJ 2259], em que intervêm formas variantes de *mor*, o que não permite a generalização do uso do hífen nesse tipo de construção. Excetuando-se a estrutura [N + mor]N, não foram registrados outros compostos NA hifenizados. Com a estrutura AN, apresenta hífen o composto *ricos-omes* [CSM 235.21] (- *rricos-homões* [CDP 17.25]), devendo-se salientar que a forma singular correspondente apresenta-se geralmente aglutinada sob a forma *ricome*.

Quanto às estruturas VN (em *guarda-cós* e *guarda-roupa*) e NN (em *ifant-abade*, *madre-donzela*), formalmente anômalas, elas também não oferecem, para o editor de texto, dificuldades de identificação como composto.

Entretanto, quanto às unidades lexicais com a estrutura NprepN, observou-se, como já foi apontado na nota 6, que em nenhum caso há a utilização do hífen: olho de boi ‘espécie de pedra (semi)preciosa’ [CEMD 377.20]; mal de pedra ‘enfermidade que resulta da formação de cálculos nas vias urinárias’ [CSM 173.10]; carta de creença ‘credencial’, ‘carta digna de fé’ [CDP 2.32], [VFDJ 3601]; *homêes das viintenas* ‘marinheiros das galés’ [CDP 14.39]; *panos de ras* [CEMD 21.23], [TCDJ 186]; *escrivam da poridade* [CDP 7.51], [VFDJ 2485]; *rabo de cavalo* [VFDJ 7402], *porteiros de maça* [VFDJ 1030]; *guia de dança* ‘aquele que comanda a dança’ [GR4 329-30], *pee de porco* ‘plectro, palheta’ [GR4 330.30]. Esse tratamento está de acordo com a concepção de que esse tipo de estrutura não constitui um composto prototípico, mas uma locução.

Resumindo, verificou-se a seguinte situação no que diz respeito ao uso/não uso do hífen nas estruturas NA, AN, NN, VN e NprepN:

Estrutura	Com hífen	Sem hífen
NA	<i>adeantado-moor</i> [CDP 16.111], <i>alcaide-moor</i> [CDP 9.42], <i>alferez-moor</i> [CDP 40.11], <i>alguazill-moor</i> [CDP 15.20], <i>beesteiro-moor</i> [CDP 20.53], <i>alcaide-mor</i> [VFDJ 6728], <i>alcaydaria-moor</i> [VFDJ 8035], <i>camareira-mor</i> [VFDJ 5295], <i>camareiro-mor</i> [VFDJ 8998], <i>capelão-mor</i> [VFDJ 7015], <i>capitão-mor</i> [VFDJ 4060], <i>contador-mor</i> [VFDJ 4336], <i>mordomo-mor</i> [VFDJ 1704], <i>porteiro-moor</i> [VFDJ 4607], <i>altar-mor</i> [TCDJ 109], <i>capela-moor</i> [TCDJ 188], <i>camareiro-mor</i> , <i>comendador-moor</i> [GR2 197], <i>Coudel-mor</i> [GR3 290], <i>estribeiro-moor</i> [GR3 157], <i>sol-posto</i> [VFDJ 373],	<i>alcaide mayor</i> [CSM 329.47], <i>altar mayor</i> [CSM 221.50], <i>monteiro maior</i> [CEMD 324.1], <i>almotace moor</i> [TT 2.5], <i>contador moor</i> [TT 2v.20], <i>mordomo moor</i> [TT 6v.8] e <i>capitam moor</i> [CC 1r.1], <i>agua rosada</i> [CSM 39.37], <i>pedras ballaises</i> [CDP 33.63], <i>guerra guerreada</i> [CDP 32.50], <i>capella mayor</i> [VFDJ 2259], <i>agoa doce</i> [VFDJ 7118], <i>Mar Ouceano</i> [VFDJ 970], <i>mal frances</i> [GR3 234.2], <i>sol posto</i> [CC 1v.8]
AN	<i>ricos-omes</i> [CSM 235.21], <i>rricos-homêes</i> [CDP 17.25]	<i>ricas donas</i> [CSM 212.13], <i>ricas donas</i> [CEMD 238.7]
NN	<i>ifânt-abade</i> [CSM 164.21], <i>madre-donzela</i> [CSM 153.50], <i>conde-prior</i> [TCDJ 73], <i>porco-pisco</i> [GR2 209]	<i>pedras marmores</i> [CSM 342.2]
VN	<i>guarda-cós</i> [CEMD 209.19], <i>guarda-roupa</i> [GR3 51.2]	<i>fura buchos</i> [CC 1v.1]
NprepN		<i>olho de boi</i> [CEMD 377.20], <i>mal de pedra</i> [CSM 173.10], <i>carta de creença</i> [CDP 2.32], [VFDJ 3601], <i>homêes das viintenas</i> [CDP 14.39], <i>panos de ras</i> [CEMD 21.23], [TCDJ 186], <i>escrivam da poridade</i> [CDP 7.51], [VFDJ 2485], <i>rabo de cavalo</i> [VFDJ 7402], <i>porteiros de maça</i> [VFDJ 1030], <i>guia de dança</i> [GR4 329.35], <i>pee de porco</i> [GR4 330.30]

Quadro 1. Estruturas composicionais com e sem hífen em edições de textos portugueses arcaicos

Complementarmente, uma observação deve ser feita em relação às formações que envolvem os advérbios *bem* e *mal* e o uso do hífen. Nas *Cantigas de Escárnio e de Maldizer*, têm-se: *ben andante* [CEMD 224.8], *ben quisto* [CEMD 287.15], *mal andante* [CEMD 103.21], *mal manhado* [CEMD 91.16], *mal parado*

[CEMD 374.14], separados e sem hífen, ao lado de *malaventura* [CEMD 98.9], *malfadado* [CEMD 386.4], *maltalhado* [CEMD 51.2], unidos, sem hífen.

Nas *Cantigas de Santa Maria*, registram-se *ben-aventurado* ‘muito feliz’ [CSM 263.4], *ben-andante* ‘feliz’ [CSM 195.200], *ben-costumado* ‘de bons costumes’ [CSM 288.38], *ben-erdado* ‘rico’ [CSM 75.26], *ben-fadado* ‘afortunado’ [CSM 135.126], *ben-fazer* ‘beneficência, caridade’ [CSM 125.4], *ben-feito* ‘benefício, favor’ [CSM 235.5], ao lado de *malandante* ‘infeliz, desgraçado’ [CSM 238.32], *malaventurado* [CSM 38.30], *malcreente* ‘descrente’ [CSM 233.33], *mal-estança* ‘inconveniência’ [CSM 9.115], *malfadado* ‘infeliz, desgraçado, infame, malvado’ [CSM 201.30], *malparado* ‘maltratado, estragado’ [CSM 5.50] (cf. *mall mao pecado* ‘infelizmente’, *s.v.* pecado [CSM 237.66]).

Essa assistemática com *bem* e *mal* pode ser observada também no novo acordo ortográfico, que, embora apregoando o uso do hífen em casos específicos, como o caso do segundo elemento iniciar-se por vogal ou *h* (e ainda *l*, no caso de *mal*, como em *mal-limpo*) – *bem-apanhado*, *mal-estar*, *bem-humorado*, *mal-habitado* – deixa aberta a possibilidade de *bem* e *mal* aglutinarem-se ou não, mesmo quando o segundo elemento iniciar-se por consoante – *bem-nascido* (cf. *malnascido*), *benfeitor* (*Escrevendo pela nova ortografia*, 2009, p. 62; 94).

A partir dessa breve análise, a seguinte questão se apresenta: a (as) sistematicidade no uso do hífen por parte dos editores de textos arcaicos reflete de fato a sua compreensão do que é um composto e o mecanismo da composição de palavras?

5 CONSIDERAÇÕES FINAIS

Buscou-se, neste trabalho, discutir sobre o uso do hífen por alguns editores selecionados de textos do português arcaico nas palavras compostas. Sabendo-se da inexistência do hífen no interior das formas compostas durante todo o período arcaico da língua portuguesa, buscou-se, ainda, a partir da observação de unidades plurilexicais registradas com (e sem) hífen, refletir sobre o conceito que os editores de texto têm de palavra composta e do mecanismo da composição de palavras.

Para desenvolver tal estudo, foi necessário, primeiramente, discorrer sobre o papel do hífen na delimitação de uma palavra composta, não apenas no português arcaico, mas também no português contemporâneo, atentando para o seu caráter convencional, uma vez que se limita ao nível da escrita. Também foi fundamental apresentar algumas considerações sobre palavra composta e mecanismo de composição de palavras, deixando evidente a dificuldade de definir esses termos tão comuns ao setor da formação de palavras, dificuldade frequentemente expressa nos estudos especializados sobre o tema. Optamos por seguir a orientação de Santos (2009) que, além de apresentar uma concepção mais abrangente do

fenômeno composicional, identifica a existência de um gradualismo estrutural e semântico perpassando todos os modelos compositivos.

A seleção do *corpus* atendeu, primeiramente, ao critério da presença de palavras compostas grafadas com hífen. Fornecem exemplos de compostos com hífen as edições das *Cantigas d'Escarnho e de Mal Dizer (CEMD)*, editados por Rodrigues Lapa; as *Cantigas de Santa Maria (CSM)*, editadas por Walter Mettmann; a *Crónica de D. Pedro (CDP)*, editada por Giuliano Macchi; *Vida e feitos d'El Rey Dom João Segundo (VFDJ)* e *Trasladação do Corpo d'El Rey Dom João o Segundo (TCDJ)*, editados por Evelina Verdelho, e o *Cancioneiro Geral de Garcia de Resende (GR)*, editado por Aida F. Dias. Também integram o *corpus* o *Tratado de Tordesilhas*, editado por Clarinda Maia, e *A Carta de Pero Vaz de Caminha*, editada por Celso Cunha, César Cambraia e Heitor Megale, edições com menor grau de intervenção dos editores. Para cada edição, atentou-se para a explicitação, nos critérios editoriais, de referência à introdução do hífen no interior das palavras compostas.

A análise dos dados permitiu-nos observar que o uso do hífen ocorre com certa sistematicidade em estruturas com clara feição de compostos, como é o caso de [N + mor], [NN] e [VN]. A estrutura [NprepN], menos prototípica em termos composicionais, talvez por ser configuracionalmente transparente, correspondendo a uma estrutura sintática da língua em que intervém uma preposição, não forneceu exemplos de formas com hífen.

Assim, retomando a questão apresentada na Introdução e no final da seção 3, acreditamos que a (as)sistematicidade no uso do hífen no registro de palavras compostas do português arcaico reflete, em alguma medida, a concepção dos editores de textos acerca do composto e do processo de composição de palavras, evidenciando a complexidade envolvida na sua definição. Nesse sentido, a opção pelo uso do hífen deve ser precedida de uma definição do que se entende por palavra composta, merecendo, este elemento, atenção por parte dos editores de textos portugueses arcaicos.

Embora o hífen não constitua um sinal de pontuação, no sentido específico da expressão, estendemos para esse elemento a seguinte reflexão de Dias *et alii*:

Quanto à pontuação original, a actualização no uso de pontos, vírgulas e outros sinais pode igualmente alterar modos de pensamento e de expressão, modificando, com escasso proveito, toda uma sintaxe que o historiador de mentalidade venha a considerar chave para a compreensão da época. (1987, p. IX, *apud* Machado Filho, 2004, p. 117)

O hífen, ou, antes, a sua ausência, nas edições de textos arcaicos, embora não esteja relacionado com a sintaxe da língua, constitui um dado importante para a reconstituição histórica da trajetória da língua portuguesa, também refletindo “modos de pensamento e de expressão”.

REFERÊNCIAS

- Cunha, Celso; Lindley Cintra, Luís F. 1985. *Nova gramática do português contemporâneo*. 2 ed., Rio de Janeiro: Nova Fronteira.
- Dicionário Houaiss da língua portuguesa. 2009. Rio de Janeiro: Objetiva.
- Dicionário Houaiss da língua portuguesa. 2004. Rio de Janeiro: Objetiva.
- Escrevendo pela nova ortografia: como usar as regras do novo acordo ortográfico da língua portuguesa. 2009. Instituto Antônio Houaiss, Coordenação e assistência de José Carlos Azeredo, 3 ed., São Paulo: Publifolha.
- Feijó, João de Moraes Madureira. 1861 [1739]. *Orthographia, ou arte de escrever e pronunciar com acerto a lingua portugueza*. 2 ed., Lisboa: Typographia Rollandiana.
- Guevara, Emiliano; Scalise, Sergio. 2009. Searching for Universals in Compounding. In: Scalise S, Magni E, Bisetto A. (Eds.). *Universals of Language Today*, Dordrecht: Springer, pp. 101-128.
- Machado Filho, Américo Venâncio Lopes. 2004. *A pontuação em manuscritos medievais portugueses*. Salvador: EDUFBA.
- Marquilhas, Rita. 1987. O acento, o hífen e as consoantes mudas nas Ortografias antigas portuguesas. In: Castro, Ivo; Duarte, Inês; Leiria, Isabel. *A demanda da ortografia portuguesa. Comentário do Acordo Ortográfico de 1986 e subsídios para a compreensão da Questão que se lhe seguiu*. Lisboa: Edições João Sá da Costa, pp. 103-116.
- Mattos e Silva, Rosa Virgínia. 2008. *O português arcaico: uma aproximação*. 2 vols. (I: Léxico e Morfologia; II: Sintaxe e Fonologia), Lisboa: Imprensa Nacional – Casa da Moeda.
- Ribeiro, Sílvia. 2006. *Compostos nominais em português: as estruturas VN, NN, NprepN e NA*. Dissertação de Mestrado em Linguística Portuguesa apresentada à Faculdade de Letras da Universidade de Coimbra.
- Ribeiro, Sílvia. Rio-Torto, Graça. 2010. Denominações compositivas de estrutura VN, NN, NprepN e NA. In: Iliescu M, Siller-Runggaldier H, Danler P. (eds.). *Actes du XXV*

Congrès International de Philologie et de Linguistique Romanes. Tome VII. Berlin, Mouton de Gruyter, pp. 477-487.

Said Ali, M. 1964. Gramática histórica da língua portuguesa. 3 ed. melhorada e aumentada de Lexeologia e Formação de Palavras e Sintaxe do Português Histórico. São Paulo: Melhoramentos.

Santos, Antonia Vieira dos. 2009. Compostos sintagmáticos nominais VN, NN, NA, AN e NprepN no português arcaico (séc. XIII-XVI). Tese de Doutorado em Letras. Programa de Pós-Graduação em Letras e Linguística, Instituto de Letras da Universidade Federal da Bahia.

Telles, Célia Marques. 2003. Que textos são oferecidos aos estudantes? Revista do GELNE, Ano 5, n. 1 e 2, pp. 21-28.

REFERÊNCIAS DOS TEXTOS DO *CORPUS*

A Carta de Pero Vaz de Caminha. Reprodução fac-similar do manuscrito com leitura justalinear de Antônio Geraldo da Cunha, César Nardelli Cambraia e Heitor Megale. São Paulo: Humanitas, 1999.

Afonso X. Cantigas de Santa Maria. Editadas por Walter Mettmann, 4 volumes. Coimbra: Acta Universitatis Conimbricensis, vol. I (1959), vol. II (1961), vol. III (1964), vol. IV (Glossário) (1972).

Cantigas d'escarnho e de mal dizer dos cancioneiros medievais galego-portugueses. Edição crítica e vocabulário do Prof. M. Rodrigues Lapa. 2 ed., Lisboa: Edições João Sá da Costa, 1988.

Crónica de Dom Pedro, de Fernão Lopes. Edizione critica, con introduzione e glossario a cura di Giuliano Macchi. Roma: Edizioni dell'Ateneo [2 ed., revista, com edição crítica, introdução, glossário e índices de Giuliano Macchi, Lisboa: Imprensa Nacional – Casa da Moeda, 2007], 1966.

Maia, Clarinda de Azevedo. O Tratado de Tordesilhas: algumas observações sobre o estado da língua portuguesa em finais do século XV. Biblos, vol. LXX, pp. 33-91, 1994.

Resende, Garcia de. Cancioneiro Geral. Fixação do texto e estudo por Aida Fernanda Dias, Lisboa: Imprensa Nacional – Casa da Moeda, 4 vols., 1990-1993 [foi selecionado volume III para a recolha das formas].

Resende, Garcia de. Livro das Obras de Garcia de Resende. Edição crítica, estudo textológico e lingüístico por Evelina Verdelho. Lisboa: Fundação Calouste Gulbenkian, 1994 [capítulos: Vida e feitos de d'el-rey Dom João Segundo e A trasladação do corpo d'el-rey Dom João o Segundo].

Recebido em: 15/11/13

Aprovado em: 02/12/13