

A GRAMATICALIZAÇÃO E A LEXICALIZAÇÃO COMO PROCESSOS HISTÓRICOS¹

MARIA JOSÉ CARVALHO
(Faculdade de Letras da Universidade de Coimbra)

ABSTRACT: This study aims to show empirically how the phenomena of grammaticalization and lexicalization can be studied from a diachronic perspective, in that they reflect historically constructed processes. Leaving aside the theoretical status of these concepts, or their terminological diversity, we emphasise here what they have in common – the fact that they are both reductive phenomena. We therefore analyse the most productive processes in the transformation of adverbs into conjunctions or discursive connectors, that is, those which, by frequency and use, and by a process which involves recategorization, lose their basic meaning (of time, mode, etc.) and come to indicate textual relationships. We also mention some simple examples of formation of autonomous lexical units (common nouns, cardinal numbers and toponyms) from complex expressions, by the (common) process of phonetic erosion.

KEYWORDS: grammaticalization, lexicalization, morphologization, subjectivization, pragmaticization.

Introdução

O fenómeno da gramaticalização tem sido abordado numa grande variedade de perspectivas, começando as divergências no tratamento deste tema com a própria terminologia usada, alguns autores preferindo “gramatização” ou ainda “gramatização” ao termo “gramaticalização”. A discordância acentua-se, contudo, quando se trata de averiguar se estamos perante um fenómeno de natureza primordialmente diacrónica ou sincrónica. Nas definições clássicas de gramaticalização, esta é definida como incidindo sobre a palavra ou o morfema, mas paralelamente a esta tradição item lexical

¹ A Autora agradece a dois revisores anónimos deste artigo, pelos comentários e sugestões feitos no sentido do seu enriquecimento, bem como pelo encorajamento que recebeu no âmbito da pesquisa sobre Gramaticalização.

> morfema, uma outra tendência mais recente tem desenvolvido trabalho muito relevante ao nível da estrutura morfológica e sintáctica de construções contextualizadas de forma específica.

Também nos últimos anos tem sido dado particular relevo à natureza histórica dos processos de gramaticalização, associando-os a um determinado tipo de mudança linguística, sujeito a certos processos e mecanismos gerais (como a transferência metafórica e metonímica, a reanálise, etc.) e caracterizado por certas consequências tais como as mudanças na gramática. Como outras mudanças linguísticas, a gramaticalização, de acordo com esta perspectiva, processa-se gradualmente, quer nos contextos linguísticos, quer nos contextos sociais. Firmemente alicerçada na perspectiva diacrónica da gramaticalização, baseia-se a assunção (já que são muito poucos os contra-exemplos citados na literatura) do princípio da não ocorrência da reversibilidade do processo, ou seja, o da *unidireccionalidade*. Grande parte da literatura sobre esta questão caracteriza o desenvolvimento do significado gramatical a partir do significado lexical em termos de *semantic bleaching* (“desemantização”, “empobrecimento semântico”), uma vez que está em causa uma perda do significado lexical e um aumento considerável do nível de abstractização do significado.² Assim, a teoria da gramaticalização tem sido, nos últimos anos, fortemente baseada na seguinte concepção de J. Paul Hopper e Elizabeth Closs Traugott: «...the change whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions and, once grammaticalized, continue to develop new grammatical functions» (Hopper and Traugott 2003²: 18). A perspectiva que, na nossa opinião, melhor tem espelhado a natureza deste processo de mudança é a que é apresentada por Elizabeth Closs Traugott, que a define como «a gradient phenomenon, whereby forms and constructions that at first express primarily concrete, lexical, and objective meanings come through repeated use in local syntactic contexts to serve increasingly abstract, pragmatic, interpersonal and speaker-based functions» (Traugott, 1995: 32). Neste sentido, a gramaticalização surge associada ao processo pragmático-semântico da subjectivização, onde «meanings become increasingly based in the speaker’s belief state/attitude toward the proposition» (Traugott, 1995: 31).

Menos sistematicamente estudado tem sido o processo de lexicalização que, tal como a gramaticalização, consiste num processo de mudança linguística que pode envolver construções sintagmáticas, formação de palavras, formações morfofonológicas, e que se caracteriza, igualmente, por uma per-

² Entendemos aqui por “desemantização”, referida, frequentemente, por “semantic bleaching” «a process whereby in specific contexts a lexical item is emptied of its lexical semantics and acquires a grammatical function». (Heine, 1993: 54). Os rótulos usados para este conceito têm sido vários, como por exemplo: “bleaching”, “desemanticization”, “semantic impoverishment”, “generalization of semantic content”, “abstraction”, “decay”, etc. Consulte-se, na obra citada, a página 89. Para uma visualização esquemática do chamado “bleaching model” consulte-se Bernd Heine, Ulrike Claudi e Friederike Hünemeyer, 1991: 109.

da de autonomia dos elementos linguísticos (nomeadamente ao nível formal, fonológico ou semântico) e num aumento da composicionalidade. Assim, numa perspectiva diacrónica, gramaticalização e lexicalização têm em comum o facto de serem processos de mudança linguística que consistem na introdução de novas formas ou significados no inventário de uma língua e nas modificações sofridas por esses itens dentro desse inventário. Em ambos os casos estamos em presença de processos contínuos, graduais e unidireccionais sujeitos aos processos de fusão. De acordo com Juan C. Moreno Cabrera, «exactly the same hierarchy that some scholars have proposed for explaining how grammaticalization takes place can be successfully applied to the opposed and in some sense contradictory process of lexicalization» (Moreno Cabrera, 1998: 224). Ora, como veremos ao longo deste artigo, os limites dos processos de gramaticalização e de lexicalização nem sempre são, na nossa opinião, fáceis de definir, apesar de reconhecermos a ausência de “semantic bleaching”³, de subjectivização, de frequência e de produtividade nos processos de lexicalização.

Neste artigo procuraremos aplicar alguns dos princípios teóricos aqui equacionados a alguns casos da história da língua portuguesa que apresentem características que os situam na gramaticalização, na lexicalização ou, em certa medida, entre ambos os processos (embora metodologicamente os tenhamos separado), porque carecendo de fronteiras claramente definidas. De facto, a análise da sua evolução histórica (desde o *source meaning* ao *target meaning*) no enquadramento teórico apresentado resolve alguns casos de impasse ou de descrição pouco clara e permite reiterar as relações de proximidade e diferenciação entre ambos os conceitos. Trata-se, em alguns casos (referimo-nos, por exemplo, aos advérbios em *-mente* ou eventualmente a alguns conectores discursivos, como *todavia* e *porem*) de um assunto que tem sido pouco desenvolvido numa perspectiva teórica ou mesmo descritivista das línguas, mas com grande repercussão no ensino das línguas e na prática da tradução.⁴

O *corpus* em que baseámos a nossa pesquisa é constituído por cerca de 153 documentos notariais originais (por nós transcritos), oriundos dos fundos do mosteiro de Santa Maria de Alcobaca (Instituto dos Arquivos Nacio-

³ Eventualmente será aqui aplicável a noção de “grammatical bleaching” ou “empobrecimento gramatical”, embora nunca tenhamos encontrado esta noção na literatura sobre lexicalização.

⁴ A este propósito, afigura-se de muito interesse o trabalho de José Luís Rodríguez (1996: 189-201) que, ao reclamar, com a persuasão científica que a matéria requer, o estabelecimento de uma norma para o galego, estuda a evolução semântica de algumas formas e/ou fórmulas conjuntivas «amiúde desconsideradas ou consideradas inadequadamente nas gramáticas galegas, em boa parte, (...), por interferência da língua do Estado» (Rodríguez, 1996: 190). Uma delas, comum à nossa pesquisa, é precisamente *todavia* (para além de *posto que* e *enquanto*), que o Autor considera um verdadeiro “falso amigo” do ponto de vista de uma linguística contrastiva luso-espanhola.

nais/Torre do Tombo), tendo sido seleccionados, fundamentalmente de acordo com critérios de carácter histórico-cronológico (sécs. XIII-XVI) e geográfico.⁵ É, de facto, nossa convicção que é na documentação notarial oriunda dos mosteiros que se poderá encontrar a génese da diversidade actual, na medida em que reflectem mais intensamente traços da língua oral da época. Estamos, no entanto, conscientes de que um estudo desta natureza não dispensaria uma comparação com o estado linguístico revelado por outras fontes documentais autênticas⁶, pesquisa que, pelas limitações inerentes ao próprio material documental remanescente, não é fácil empreender.

2. “Gramaticalização”⁷

2.1. O advérbio *embora*

Por vezes, a redução fonética envolvida na gramaticalização passou por um processo de erosão que aproximou a(s) forma(s) em causa daquela(s) que vigora(m) actualmente. Contudo, no português medieval, esses processos de erosão fonética nem sempre foram acompanhados de uma mudança de categoria sintáctico-semântica. É o que se passa com o advérbio *embora*, presente no exemplo seguinte:

- (1) «E começarã de fazer a primeira paga per dia de lam Joam Baptifla que ora *embora* vem de b^c xxix» (1528 MA 147)

De facto, *embora* pode parafrasear-se pelo sintagma ‘em boa hora’, que parece não ter ainda o valor de concessão, no século XVI.⁸ Assim, não obs-

⁵ A identificação do documento é feita através da indicação do ano, seguida da sigla do local onde foi redigido e do número que o acompanha na nossa edição (Carvalho, 2006). São as seguintes as siglas usadas e respectivos locais: Alc (Alcobaça), Alf (Alfeizerão), Alj (Aljubarrota), Alp (Alpedriz), Alv (Alvorninha), AM (Aldeia do Mosteiro), Cel (Cela), Evo (Évora de Alcobaça), MA (Mosteiro de Alcobaça), Mai (Maiorga), Ped (Pederneira), PP (Póvoa das Paredes), Sal (Salir do Mato), SC (Santa Catarina), SM (S. Martinho) e Tur (Turquel).

⁶ Os documentos analisados são todos autênticos, ou seja, não resultam de cópias nem contêm cartas insertas atribuíveis a outras mãos. Seria interessante comparar a situação dos documentos de natureza notarial com os textos literários, embora a tradição textual que caracteriza estes últimos (cujos contornos temporais muitas vezes desconhecemos) se torne desconfortável para o historiador da língua, que lida com cronologias precisas.

⁷ O termo “gramaticalização” aqui utilizado distancia-se, em alguns exemplos, da noção de “gramaticalização” de filiação traugottiana, na medida em que nem sempre está em causa um processo de *subjectivização*. Na verdade, o processo envolvido em alguns exemplos é, muitas vezes, um fenómeno de “morfologização”.

⁸ Segundo José Pinto de Lima, «o sintagma *em boa hora* era empregue nesta altura [séculos XV-XVI] para qualificar a acção ou acontecimento expresso pela proposição como sendo de bom augúrio, ou como ocorrendo em hora propícia ou favorável» (Lima, 1997: 646).

tante a erosão fonética que provocou a aglutinação, a sua evolução semântico-pragmática não acompanhou a «tendência para passar a grafar a expressão como uma única palavra» (Lima, 1997: 647).

2.2. Da construção circunstancial *toda uia/todauia* ‘sempre, constantemente’ ao conector adversativo⁹

Uma outra forma que importa destacar é *todauia*, cujos elementos compositivos surgem unidos em 61,5% dos casos. Contudo, ao longo de todo o *corpus* convivem as formas com e sem união dos seus elementos. Assim, por volta do terceiro quartel do século XIV surgem *todauia* (1375 MA 48) e *toda uja* (1372 MA 47); o mesmo acontece na década de 20 do século XV: *toda vja* (1422 MA 82) e *todauja* (1426 MA 85). Esta convivência verifica-se ainda em finais do mesmo século: *todavia* (1489 MA 130) e *toda via* (1490 MA 131). Convém, sobretudo, salientar os aspectos relativos à elevação desta construção circunstancial à categoria de conector adversativo, independentemente de se encontrarem ou não unidos os seus elementos constitutivos. Assim, no século XIV *todauia* mantém ainda o seu sentido etimológico (‘sempre’, ‘constantemente’), como demonstram os exemplos:

- (1) «e goardardes *toda uja* os ditos gaados que fe nõ chegẽ aos comoros das abertas nõ pafem per elas» (1372 MA 47)

Por vezes, *todauia* constitui um reforço adverbial de *sempre*, como no exemplo seguinte, que apresenta este advérbio seguido de *todauia*:

- (2) «e andarem *fenpre todauia* em hũa pefoa como dicto he» (1375 MA 48)

Fornecendo abundante exemplificação na prosa literária galega do século XIV, José Luíz Rodríguez admite, contudo, que «difícil é, assim mesmo, explicitar a evolução semântica que conduz, ao longo do séc. XIV, do valor adverbial ao conjuncional, a qual parece passar por uma ampliação da esfera temporal (‘em todo momento’) à modal (‘de todas as formas’), que conle-va já um certo carácter adversativo, sobretudo quando reforçado, por contraste, com o enunciado na oração anterior» (Rodríguez, 1996: 194).

Este autor chama, no entanto, a atenção para uma dezena de exemplos extraídos de documentação trecentista do convento de Santa Maria de Belvis, em Santiago (1336-1398), onde surge uma fórmula recorrente em que se insere a forma *todavia*, com um significado adversativo «atemperado», equivalente aproximadamente a ‘em todo caso, de todas as formas’ (Rodríguez,

⁹ Parecida evolução se deu em francês e em italiano: *toutefois* (fr.) e *tuttavia* (it.), ‘não obstante’.

1996: 195-196). No nosso *corpus*, da mesma natureza, surgem igualmente exemplos congêneres, mas apenas no século XV, o que nos leva a concluir que, na área geográfica sob análise (o Centro-litoral de Portugal), parece ter sido nesse século que o processo de *recategorização* se verificou. Este processo deverá ter sido gradual (como, aliás, o são todos os fenómenos de gramaticalização), tendo seguido o modelo: “quer se realize A, quer se realize B → faça-se *todavia* C”. De facto, embora o sentido etimológico (‘em todas as circunstâncias’) se preserve, aflora já a ideia de contraste, de “counter-expectation”, constituindo esta uma questão de ponto de vista e, portanto, um exemplo de subjectividade (Traugott, 1999: 178-179), como veremos adiante. É o que acontece nos seguintes exemplos das cláusulas formulares, que atravessam o século XV:

- (3) «*e a dicta pena paguada ou ñ **todauja** a dicta avença feer firme antre elles e estaujll pera ffenpre, como dicto he*» (1402 MA 67)
- (4) «*E pagada a dita pena ou ñ, **todauja** a dita cõuêça ficar firme*» (1438 Ped 95)
- (5) «*E levada a dita pena ou ñ prometerom os ditos dom abbade e fobprior e samchriftão e monjes e convento (...) de **toda via** este contrauto de treslauçam feer firme e valioffo pera todo fenpre*» (1490 MA 131)

É interessante constatar que na primeira metade do século XV, *todauia* surge já na acepção de ‘em todo o caso’, literalmente parafraseável, aliás, por ‘sempre’, mas contextualmente veiculando a noção de contraste e, portanto, com o mesmo sentido adversativo de ‘contudo’¹⁰:

- (6) «*djzia o dicto frey Johã ao dicto Afomfê Anes que quiria escaybar a fua meatade cõ a dicta pytançaria por outra caã tam boa no dicto logo, ou fe almoadafe pera fe vender (a) quẽ por ella mais defe ou fe alugafe de permeo a quẽ por ella mais defe e partifem o aluguer de permeo, pois que fe ñ podia partjr cõ proll da dicta caã. E que elle, dicto jujz, fezera pergunta ao dicto Afonfê Anes que por o que dizia ao que lhe o dicto frey Johanne dizia, e que o dicto Afonfê Anes *difera que ell ñ quiria e que **todauja** que fe partife* pera ell auer o feu quinhã» (1402 Ped 68).*

Todauia pode ainda ter o sentido de ‘ainda’, ‘ademais’ como no seguinte exemplo extraído de um documento de S. Martinho de 1442:

¹⁰ Os exemplos da primeira metade do século XV apresentados por Clarinda de Azevedo Maia têm ainda acepção etimológica, excepto o que se situa num documento de 1450 em que «a forma *toda via* é já utilizada claramente como conjunção, de valor equivalente a ‘contudo’» (Maia, 1997: 882).

- (7) «e dife aos sobre dictos que elles amdauã cõ dous batees e rredes e companhia, pescãdo no porto do dicto logo de fã Martinho e que lhe rrequeria da parte do ōenhor dom Esteuã d'Agjar, abade do dicto mosteiro, que lhe defem a dizima da sardinha e doutro quall quer pescado que ao dicto porto trouuesem, (...) e que *lhes rrequeria que todauja lhe defẽ* a dizima da sardinha que ao dicto porto trouxerõ e defcarõ» (1442 SM 97).

Os exemplos extraídos situam-se em documentos redigidos nos coutos que, pela menor rigidez formal, proporcionam um afloramento mais intenso de traços da língua oral.

Foi assim que, de construção circunstancial com o significado etimológico de 'sempre', *todauia* se recategorizou em conjunção adversativa, constituindo, assim, um caso de subjectivização, na acepção traugottiana: «When a speaker expresses counter-expectation, he or she expresses beliefs or points of view contrary to his or her own or the interlocutor's expectations regarding the states of affairs under discussion» (Traugott, 1999: 178).

Assim, pode dizer-se que, a partir do século XV, os falantes que actualizassem «faça-se *todavia* X» não só exprimiriam que X deveria fazer-se sempre, em todas as circunstâncias, mas conversacionalmente implicavam que X deveria fazer-se, mesmo que se verificasse Y, ainda que Y constituísse um acontecimento de relevo. Esta implicatura – por força da frequência de ocorrência – tornou-se convencional, daí a noção de adversativa ter-se tornado parte da semântica de *todavia*, sendo a ideia de contraste pragmaticamente mais relevante que a de frequência temporal.

Curiosamente, também a partir do século XV, *todauia* começa a surgir nos documentos redigidos no mosteiro e com uma certa rigidez formular, como reforço dos conectores adversativos *mas* ou *salvo*. Esta necessidade de reforço atesta que a nova acepção de *todauia*, já distante do significado literal latino, não era ainda pacificamente aceite nos registos mais formais:

- (8) «e cõtra esto nõ ualer ley de liuridõẽ nõ graça nõ derecto nõ priuilegio de fidalgo nõ doutra qual quer maneira que feia (...), *mais toda uia* perante o dicto nõsõ ouujdor rresponderdes e ouujrdes sentença em o dicto nõsõ moefteiro, como dicto he» (1403 MA 69).
- (9) «E pera esto uos nõ valham os priuilegjos nõsõ dictos todos nõ cada hũu delles, *majs todauja* rresponderdes perante ell em o dicto nõsõ moefteiro e hi ouuirdes sentença» (1419 MA 79).
- (10) «que de todo uos nõ pofades ajudar cõtra esto, *saluo* pagar *toda vja* cõprindossẽ todo pella guifa que dicto he» (1422 MA 82).
- (11) «que lhe nõ balha nõ fẽ poffa delo ajudar, *saluo todauja* rresponder e pagar como dicto he» (1426 MA 85).

- (12) «que lhe nõ ualha, *saluo todauja* rresponder e pagar como dicto he» (1428 MA 87).
- (13) «e flê todo nom conprir que uos nõ valha *mays todauja* em todo conprirdes este contrauto fllegũdo em elle he contiudo» (1452 MA 106).
- (14) «e flê todo nõ conprir que uos nõ valha, *mays todauja* em todo conprirdes e mãterdes este enprazamêto» (1453 MA 107)
- (15) «*saluo todavia* conprir e mãteer e pagar e ffazer todo ho conthiudo em este stromêto» (1489 MA 130).

2.3. *Porende/porem*: do valor conclusivo-explicativo ao conector adversativo

Analisemos, agora, a história do conector adversativo *porem*. É sabido que *porém* evoluiu do sintagma circunstancial latino constituído por preposição e elemento pronominal (< PERİNDE), tendo como equivalente semântico “por causa disso”, “por esse motivo”, “por isso”. É, aliás, com esse valor conclusivo-explicativo que *porem* (forma apocopada de *porende*, extremamente arcaico) aparece ao longo de todo o período medieval. Os exemplos mais antigos documentados no *corpus* em análise são os seguintes:

- (1) «fizemos fazer *porêde* dũa cartas» (1291 Alc 3)
- (2) «e do preço apref de uof nêhũa coufa ficou pera dar. *Porê*, aiadef uof a dauãdita herdade pera todo sempre possfloydoyra (...)» (1300 Alj 8)
- (3) «forõ meter furtiuilmête marcos e deuyfões pela mata do abade e da orden d’Alcobaça fen feu mãdado e fen feu procurador, e pedía *porê* que os matafen» (1329 Evo 22)
- (4) «E deziã que por que Johã foarez e Martin Anes fe forõ meter e laurar no dicto logo, que eles pera ueerê fe lhjs entrauan naquel logo en que nõ auyã d’auer outros lauradores, que *porê* per encoberta meterõ os dictos marcos e deuyfões» (1329 Evo 22)

Não obstante Rosa Virgínia Mattos e Silva considerar que só no século XVI «se difunde o valor adversativo de *porém* (...) que se mantém em uso até os nossos dias, em detrimento de seu valor etimológico» (Silva, 1984: 130), os documentos notariais sob análise revelam, contudo, alguma precocidade no seu uso com valor adversativo. Os contextos que parecem ter, de facto, propiciado a transição semântica foram as frases negativas, como já referiu Said Ali (Ali, 1964: 187). Assim, por exemplo, no documento 1329 Evo 22, *porê* com valor explicativo-conclusivo coabita com *porêde*, já com valor adversativo (apesar de formalmente mais arcaica), sempre em frases negativas. Curiosamente, mantém-se ainda o significado primitivo de *poren-*

de ('por isso'), mas começa a desenvolver-se um novo sentido que vai constituir a pragmática de *porem*, sentido esse que deverá ter nascido aliado a frases negativas, como já se disse. Assim, já na primeira metade do século XIV, um falante que proferisse «não ...*porem* X», não só exprimiria que X é consequência de alguma acção anteriormente mencionada discursivamente como implicaria que X, a ocorrer, contraria a perspectiva/ponto de vista do sujeito sintáctico:

- (5) «E os dictos prefos cõfesarõ dante mjm en juízo que meterõ os dictos marcos e deuyldes mais deziã que nõ deuyã *porẽde* de morrer» (1329 Evo 22)
- (6) «Deziã outrofy que, como quer que o dicto logo en que chãtarõ os dictos marcos nõ ffofe feu proprio mais da ordẽ, que nõ leyxauã *porẽde* a uer directo en efẽ logo» (1329 Evo 22)
- (7) «e afy deziã os dictos prefos que per todas eftas rrazões e por cada hũa delas que nõ deuyã *porẽde* a morrer» (1329 Evo 22)
- (8) «julgey que os dictos prefos nõ deuã *porẽde* morrer» (1329 Evo 22)

Os exemplos arrolados matizam, assim, a hipótese do «carácter estável do item *porém*, como conclusivo-explicativo», «até o século XIV, antes de suas duas últimas décadas» (Silva, 1984: 142).

A partir de 1375 e até cerca da década de 80 do século XV *porem* é, normalmente, usado como conector discursivo com valor explicativo-conclusivo, embora os sentidos sejam já ambíguos, em muitos dos casos:

- (9) «*Por* ã, dou o dicto moefteiro pera fẽpre por quite e liure da dicta contia, das quaees coufas o dicto celareiro pedjo hũu estromẽto» (1375 MA 49)
- (10) «E *porem* o dicto Senhor dom abbade e conuẽto e os dictos concelhos mãdarõ feelar esta carta da dicta auẽẽça dos feus feelos em pendiente» (1391 MA 59)
- (11) «E que por que a dicta cafa era tam pequena que fe nõ podia partjr ã guifa que fe em ella podefem fazer duas moradas, que *porẽ* djzia o dicto frey Johã ao dicto Afomfe Anes que quiria escaybar a sua meatade cõ a dicta pytançaria por outra cafa tam boa no dicto logo» (1402 Ped 68)
- (12) «E, *porẽ*, o dicto Fernãde Anes diffe que frõtãua e rrequeria (...)» (1412 Ped 74)
- (13) «e que *por em* pidia aa mjm que, per mjnha sentença julgãdo (...)» (1421 Evo 80)

- (14) «*Por em* uos mândo que cõprades e façades cõprir e aguardar o dicto meu juízo» (1421 Evo 80)
- (15) «de que elle eſtaua bem ſanhoſo, e que *porem* lhe pedhiã perdom por o amor de Deus» (1436 Alf 93)
- (16) «e o dicto ſenhor lhe mādou *porem* lleuar ſeu mōtado derreito» (1436 Alf 93)
- (17) «fuy bem pagado e ſatisfecto que nhuã coufa me nõ ficou por pagar. *Por em*, dou dello por quity e ljuri deſte dia pera todo lēpre» (1443 Alf 99)
- (18) «e que *porem* ell ffazia pergunta ao dicto Afomfo Lourenço, que hi no prefente eſtaua» (1444 Alv 100)
- (19) «*Por em* uos rrequero da parte d’el Rej que facaaes conprir e guardar a dicta fētença» (1451 MA 105)
- (20) «*Porem*, enprazamos todos juntamente a uos» (1452 MA 106)
- (21) «E que, *porem*, querendo elle, dito dom abade, auer compayxom com os ditos conçelhos que allĩ aujam de ferujr, que lhe prazia dar algum tanto de mantjamento» (1456 MA 109)
- (22) «*Porem* uos rrequero da parte d’el Rej que façaaes conprir a dicta fentēça ſegũdo per vos he julgado» (1460 MA 113)
- (23) «E, *porēm*, mādārõm e outorguarõm (...)» (1462 Mai 114)
- (24) «*Porem*, quero e outorgo (...)» (1469 Cel 118)
- (25) «e que, *porem*, elle llob prioll e cõuēto, (...) differam que elle ãprazauã, como logo de fecto ãprazaram» (1485 MA 128)

É a partir da década de 80 do século XV que *porem* adquire um valor semântico adversativo, no nosso *corpus*¹¹:

- (26) «poſſã tomar poſſẽ da dita almoinha e fazer della como coufa ſua, ficando, *poren*, os lobrebitos a cõpoer a ſuas cuſtas todo denjficamento que achado for na dita almojnha» (1482 MA 125)

¹¹ Curiosamente, a conjunção *pero* tem sempre, no nosso *corpus*, um valor adversativo, documentando-se a última ocorrência num documento de 1451. De facto, «havendo um novo item, *porém*, para a expressão da adversativa, não faria falta o *pero* adversativo: *porém* ocuparia a «casa vazia de *pero*» (Silva, 1984:149). Alguns dados interessantes sobre a polissemia de *pero* num *corpus* literário trecentista encontram-se neste artigo da Autora citada.

- (27) «E fendo nos afy rrefetoydos, o djto Fernã Carualho outra vez nos torna a forçar, tomãdo e roubãdo os djzjmos da djta egreya. *Porem*, lhe rrequeremos da parte da fanta madre egreya que cefe de nos majs fforçar (...)» (1491 Alj 132)
- (28) «e hora nos he djto que voos pproçedees contra nos fem nos mofttrardes per honde nem que autorydade tendes. *Porẽm*, nos, cõff-yãdo na djta apellaçãm, vos rrequeremos da parte do fãto padre que nos dees o trelado per eſcripto da autorydade que tendes» (1491 Alj 133)
- (29) «pello dicto Johã Afomlo, rreeo, ffoy dicto que o dicto juiz lhe era folpeyto. *Porem*, que ele, como ffora de nõ feer feu juiz, dyzya que nõ querja demãda com ho dicto moefteyro» (1496 Sal 135)
- (30) «E nã has querendo, que emtã o poſam fazer com licença do dito moefteiro, e doutra maneira nam, nom ſendo *porẽ* a nhũa peſoa poderofã e das em derreito defellãs» (1522 MA 144)
- (31) «E não as querendo, que então o poſam fazer e doutra maneira nã, nã ſendo *porẽ* a nhũa peſoa poderofã» (1528 MA 147)

2.4. Os advérbios em *-mente*

Semanticamente, a forma de Ablativo latino MENTE perdeu o seu carácter de restrição a estados psicológicos. O então item lexical *mente* evoluiu para um afixo que se acrescentou a um adjetivo, tendo-se formado, assim, um advérbio. Este caso é, assim, tipicamente, um caso de “morfologização”. Vestígios da sua origem lexical, convivendo já com o seu novo estatuto de morfema encontram-se no seguinte exemplo trecentista do nosso *corpus*:

- (1) «ca lhy fezera amor e *grã mẽte* fora ſã voõtade e *mẽte*¹² ouue de ueer o da ordẽ» (1326 MA 19)

Geralmente, *-mente* encontra-se ligado ao adjetivo, mas há alguns casos em que surge separado. Para além do exemplo supra citado, registam-se 10 casos de separação, todos registados em documentos posteriores à década de 30 do século XV. No total, não excedem os 4% do total. São os seguintes¹³:

¹² A expressão «*aver em mente*» ‘pensar em’, ‘lembrar-se de’ encontra-se largamente documentada no português medieval. Veja-se, por exemplo, Mettmann, (1972): s. u. “mente”: «E en *mente* non avia do que prometera» (117.16). O sentido desta expressão, no exemplo por nós citado, parece ser ‘ter intenção de’.

¹³ Com união de palavras, ao longo do leque cronológico abrangido por este estudo, contam-se 41 ocorrências do advérbio *cõpridamente*, 68 de *fielmente*, 26 de *liuremente* e 15 de *primeiramente* (e respectivas variantes gráficas).

cõprida mēte (1438 Ped 95), *conprida mēte* (1462 Alj 115), *clara mēte* (1565 Alc 153), *cõthynuada mēte* (1434 SC 91), *fiell mēte* (1479 MA 124), *liure mente* (1490 MA 131), *perfeita mēte* (1438 Ped 95), *primeira mēte* (1459 MA 110), *primeyra mēte* (1519 MA 142) e *verdadeira mēte* (1515 SM 141).

Estes dados de natureza cronológica indiciam, na nossa opinião, o surgimento de uma clara consciência da composicionalidade morfemática e, portanto, um interessante caso de consciência (meta)linguística.

2.5. O advérbio *peruentura*

Analisemos agora um caso de união de palavras que, após ter apresentado erosão fonética, resultou num fenómeno de gramaticalização. Referimo-nos a *peruentura*, cujas variantes ducentistas encontradas neste *corpus* são as seguintes:

«e fe pela uētura» (1289 MA 1), «e fe per auētura» (1291 Alc 3) e «e ffe pela uētura» (1297 Alc 5)

A tendência dominante é «per uētura/uentura», ocupando a variante com união de palavras a percentagem de 26,6%. Situam-se cronologicamente em finais do século XIV e finais do seguinte:

peruētura (1397 MA 63 e 1491 Alj 133) e *peruentura* (1399 MA 66 e 1491 Alj 133)

3. “Lexicalização”

O fenómeno de lexicalização mais interessante do nosso *corpus* diz respeito à aglutinação (acompanhada de erosão fonética) verificada em *fidalgo*, já em finais do século XIII: *fidalgo* (1297 Alc 5). Nos documentos do século XIV é a antiga variante que se documenta, ou seja, aquela que mantém os seus elementos constitutivos:

filha d’algo (1324 Alc 18), *filhos d’algo* (1321 Alc 17; 1397 MA 63) e *ffilhos d’algo* (1399 MA 66)

No século XV, nos três documentos que registam esta forma, apenas o redigido em Alvorninha apresenta a antiga variante: *ffylhos d’algo* (1450 Alv 104), conhecendo os outros dois a variante moderna: *fidalgo* (1403 MA 69 e 1490 MA 131).

Por outro lado, é igualmente surpreendedora a precocidade revelada em *ricofoméés* (1291 Alc 3) relativamente à justaposição dos seus elementos lexicais constitutivos, uma vez que os exemplos congéneres apresentam nos séculos seguintes a sua separação gráfica:

hufo fruuto (1388 MA 57, 2 v. e 1399 MA 65, 2 v.), *mall querença* (1402 MA 67), *sobre juíz* (1328 Alj 21), *sobre juízes* (1328 Alj 21), *sobre juizes* (1391 MA 59), etc.

Uma vez que o primeiro dos elementos de *ricofoméés* termina em fonema sibilante, este exemplo trecentista evidencia, assim, o papel da fonologia nos casos de gramaticalização e de lexicalização.

Um outro aspecto que merece algum relevo é a lexicalização de numerais cardinais constituídos pela unidade *dez*. A sua união a outro numeral é, normalmente, efectuada através da conjunção de coordenação *e*, processo mais frequente até cerca de meados do século XV:

dez e noue (1340 Ped 29; 1460 MA 113), *dez e oyto* (1375 MA 49; 1380 Alv 52; 1405 MA 70; 1408 MA 71; 1477 MA 121), *dez e sete* (1412 Ped 74), *dez e seys* (1412 Ped 74), *dez e fejs* (1372 MA 47), *dez e fete* (1416 MA 78; 1433 Ped 90), *dez e ffete* (1485 MA 128), *dez e ffete* (1402 MA 67), *dez e noue* (1435 Alj 92), *dez e octo* (1297 Alc 5), *dez e noue* (1435 Alj 92).

No último quartel do século XIV, encontra-se o primeiro indício da tendência para a lexicalização dos dois elementos, sem a conjunção de coordenação *e* como elemento de ligação: *dez oyto* (1377 Alv 50). Por outro lado, essa conjunção mantém-se na última década desse século, ligando graficamente os dois elementos: *dezezex* (1391 MA 59). Só depois de terminado o primeiro quartel do século XV, a conjunção começa a desaparecer, dando origem a flutuações coevas do tipo: *dez ojto* (1429 MA 88) e *dezooyto* (1430 Cós 89). A partir da década de 50 do século XV, *e* começa a sofrer erosão fonética, sendo substituída pelo elemento vocálico *a*¹⁴:

dez a ffeys (1452 MA 106), *dezanoue* (1496 Sal 135), *dezafejs* (1502 MA 137), *dezafete* (1467 Mai 117; 1491 Alj 133), *dezooyto* (1526 Ped 145) e *dzasete* (1526 Ped 145).

Um caso diferente, uma vez que evoluiu em sentido inverso, é o da lexicalização de numerais cardinais com o nome *cento* como elemento constitutivo. No total, cerca de 69% das ocorrências apresentam a união do numeral cardinal ao nome *cento*, tal como acontece actualmente. Aliás, nos documentos datados até cerca de 1425, é essa a tendência exclusiva:

quatroçentos (1385 MA 54; 1386 MA 55), *quatrocentos* (1426 MA 85), *quatrocentos* (1362 MA 44; 1386 MA 56; 1399 MA 65; 1402 MA 67; 1405 MA 70, 2 v.), *quatroçentos* (1372 MA 47; 1377 Alv 50; 1380 Alv 52; 1392 MA 60; 1402 Ped 68; 1419 MA 79), *quatroçêntos* (1396 Ped 62), *quatroçêtos* (1362

¹⁴ Segundo José Joaquim Nunes, a língua portuguesa, como as demais línguas românicas, preferiu as expressões latinas analíticas às sintéticas e «assim, exprimiu aqueles números [de 16 a 19] por: *dez e seis*, *dez e sete*, *dez e oito* e *dez e nove*; mais tarde, porém, substituiu a conjunção copulativa *e* pela proposição [sic] *a*, dizendo hoje *dezasseis*, *dezassete*, *dezoito* e *dezanove* (...)). E acrescenta em nota que Bourciez explica este *a* pela conjunção copulativa *ac*: «a mesma partícula entra na formação popular de *vinte* por diante, como se vê em *vint'a um*, *vint'a dois*, (...)». (Nunes, 1989: 210 e nota 2 da mesma página). Numerosas abonações dos diversos tipos de variantes podem ser encontradas em Lorenzo, (1977) (s. u. “dez”).

MA 43; 1363 MA 45; 1375 MA 48; 1388 MA 57), *quatroçetos* (1366 MA 46), *quatroçêtos* (1379 Alc 51), *quatroçêtof* (1412 Ped 74; 1415 Ped 77), etc.

A partir do segundo quartel do século XV, a união e a separação dos dois elementos constitutivos começam a competir, reflectindo, assim, uma profunda consciência relativamente à composicionalidade lexical deste produto. Apresentam-se a seguir os exemplos situados entre 1425 e 1527, e que representam, nesse período cronológico, 50% do total¹⁵:

quatro çentos (1435 Alj 92), *quatro çentos* (1447 Alj 101; 1448 Alj 103; 1452 MA 106; 1453 MA 107, 2 v.), *quatro centos* (1448 Alj 103), *quatro çêtos* (1430 Cós 89; 1434 SC 91), *quatro çentos* (1440 MA 96; 1442 SM 97), *quatro çêtos* (1438 Ped 95), *quatro çentos* (1428 Alj 86), *quatro çentos* (1455 MA 108), *seis centos* (1527 MA 146) e *fete çentos* (1490 MA 131).

Convém referir que no documento quinhentista 1527 MA 146 convivem *seis centos* e *seisçentos* (2 v.).

Por último, importa fazer uma pequena alusão a alguns aspectos da lexicalização onomástica. Assim, por exemplo, o actual topónimo *Alentejo* não tinha ainda sofrido o processo de cristalização onomástica que consistiu na fusão do advérbio *alem* com o nome de rio *Tejo*, e que deu origem à designação de província. Senão vejamos:

- (1) «noſſo celareyro e procurador d'áalem Teio» (1298 Alc 6)
- (2) «deffa comarca d'aalẽ Teio» (1397 MA 64)

Por outro lado, se em 1304 encontramos o topónimo *Mõte Mayor o nouo* (1304 Alc 9), na segunda metade desse século este topónimo apresenta já uma evolução fonética: *Mõte Moor* (1372 MA 47). Curiosamente, num documento do século XVI convivem as variantes *Mõte Moor* e *Monte Mor* (1515 SM 141), esta última apresentando a crase das duas vogais idênticas.

O exemplo mais interessante é, sem dúvida, o de *Aljubarrota*. Sem nos determos nas diversas configurações gráficas deste topónimo (na fase mais arcaica *Aliumarrota/Aljumarrota*), centrar-nos-emos nos casos em que os dois elementos constitutivos (N + Adj.) se encontram separados. Mais uma vez, ao contrário do que seria de esperar, contam-se apenas 4 casos nessas condições, constituindo, assim, cerca de 9% do total de ocorrências:

Alguba Rota (1421 Evo 80) e *Aljuba Rota* (1428 Alj 86; 1448 Alj 103; 1472 TC 120)

¹⁵ Vejam-se a seguir alguns exemplos com união de palavras nesse escopo cronológico: *oytoçentos* (1495 MA 134), *quatroçentos* (1482 MA 125, 2 v.), *quatrocentos* (1444 Alv 100; 1448 Ped 102; 1456 MA 109), *quatroçêtos* (1442 MA 98), *quatroçêtos* (1433 Ped 90), *feteçentos* (1428 MA 87), *feteçêtas* (1442 MA 98, 2 v.), etc.

Como se pode verificar, os exemplos são posteriores a 1420 e situam-se em documentos redigidos nos coutos. Esta constatação, para além de levar a considerar a existência de uma forte consciência metalinguística no que se refere à composicionalidade lexical das palavras, mostra que nos meios rurais a erosão fonética que levou à cristalização toponímica das duas unidades se processava de forma mais lenta.

Conclusões

Tem sido referido (Heine, 2002: 83) que em muitos estudos de gramaticalização o procedimento adoptado é contrastar uma unidade ou construção formal e semanticamente mais precoce do ponto de vista histórico (A) com uma forma mais tardia (B) e, descrevendo as diferenças entre as duas, interpretar o que aconteceu no percurso de A até B. Negligencia-se, assim, muitas vezes, a natureza gradual dos processos de mudança e dos seus estádios intermédios:

What is frequently underrated in such studies is that the process from A to B is a *continuous* one, involving a multitude of intermediate stages, and that without a more comprehensive understanding of the nature of these intermediate stages, generalizations on the overall process must remain preliminary.¹⁶

O carácter gradual dos processos de gramaticalização analisados leva, inequivocamente, à confirmação da noção de “estádios de gramaticalização” (“*stages of grammaticalization*”) e de “cadeias de gramaticalização” (“*grammaticalization chains*”), já referidas por este autor (Heine, 1993: 120):

Since with the development of a new stage, previous stages do not disappear from the language, the result is a series of overlapping grammaticalization stages that co-exist in the synchronic structure of that language in the form of grammaticalization chains.

Por outro lado, confirma a noção de “layering” proposta por Paul J. Hopper (1991: 23):

This formal diversity comes about because when a form or set of forms emerges in a functional domain, it does not immediately (and may never) replace an already existing set of functionally equivalent forms, but rather the two sets of forms co-exist. They may be specialized for particular lexical items, particular classes of constructions, or sociolinguistic registers; they may have slightly different meanings, or simply be recognized as “stylistic” alternatives.

¹⁶ O sublinhado é nosso.

As diferentes cronologias e os diferentes matizes semânticos observados na comparação da evolução de alguns itens (*todavia*¹⁷, por exemplo) nas línguas e modalidades linguísticas peninsulares (Rodríguez, 1996: 189--201) aconselham à aplicação do conceito de *estádio de gramaticalização* não apenas ao estudo de uma língua em particular, mas igualmente à evolução e descrição de famílias de línguas.¹⁸

Tentámos mostrar, ao longo desta análise, que o estudo dos contextos que figuram na evolução gramatical são condição *sine qua non* para compreender como determinadas formas com significação lexical dão lugar a formas com novos significados (gramaticais).¹⁹ Por outro lado, alguns fenómenos de recategorização estudados, revelando uma mudança gradual no significado dos itens em causa, são nitidamente instâncias de subjectivização, no sentido em que esta é entendida por Elizabeth Traugott, uma vez que aquilo que é dito é gradualmente substituído pelo aqui-e-agora do sujeito que fala/escreve, considerado o “local” da relevância. Assim, uma vez que «meanings become more speaker-based», esse sujeito implícita conversacionalmente significados que não estão codificados linguisticamente. O seu uso repetido em contextos sintácticos localizados leva a que essas implicaturas se convencionalizem.

Por outro lado, alguns fenómenos, por ocorrerem em documentos de áreas periféricas ao mosteiro (de carácter rural), poderão considerar-se sintomas da necessidade de relacionar os estudos sobre gramaticalização com pesquisas de natureza sociolinguística. De facto, «unfortunately, a fuller integration of sociolinguistic and developmental research with research on grammaticalization still remains to be worked out» (Hopper e Traugott, 1993: 30). Para além de parâmetros de natureza social, o estilo textual parece condicionar a escolha das variantes, pelo que seria desejável estender este tipo de pesquisa a um *corpus* mais alargado, transcrito com critérios linguísticos rigorosos, que incorporasse outros géneros textuais autênticos: «specialization does not necessarily entail the elimination of alternatives, but may

¹⁷ De acordo com José Luís Rodríguez (1996: 200-201), *todavía* (reformulado como advérbio de tempo segundo o modelo espanhol, sentido popularmente como exterior à língua e caracterizador de bilingues de língua materna castelhana) é, como tantas outras, uma forma tradicional da Galiza e, ao mesmo tempo, actuante no eixo Portugal-Brasil, pelo que não deveria ceder a *sen embargo*, neologismo de procedência espanhola favorecido pelos canonizadores oficiais.

¹⁸ Relativamente ao comportamento semântico de *todavía*, é interessante e pertinente considerar que o rio Minho «nom chegou a ser fronteira, e, menos ainda, lingüística» (Rodríguez, 1996: 201).

¹⁹ Torna-se, assim, crucial na mudança semântica a noção de “bridging contexts”, que na literatura de filiação griceana tem sido descrito como “inferências” e “implicaturas”. Segundo Heine (2002, 84-85) os chamados “bridging contexts” podem dar origem a significados gramaticais convencionais, não directamente mas através dos “switch contexts”, onde o significado de chegada (*target meaning*) é incompatível com o significado de origem (*source meaning*).

be manifested simply as textual preferences, conditioned by semantic types, sociolinguistic contexts, discourse genres, and other factors» (Hopper e Traugott, 1993:114).

Os fenómenos de lexicalização, por seu turno, parecem ter revelado mais inércia nos documentos das áreas periféricas. Este facto, denotando uma forte consciência da composicionalidade das palavras, por parte dos escribas, poderá falar a favor de uma maior resistência à erosão relativamente aos fenómenos de redução lexical, nas classes menos cultas (eventualmente, por ultra-correcção), mas será necessário um trabalho mais sistemático nesta área para se avançar nas generalizações.

Referências

- Ali, Said (1964). *Gramática histórica da língua portuguesa*. 3.^a edição revista e aumentada. S. Paulo: Edições Melhoramentos.
- Brinton, Laurel J. and Elizabeth Closs Traugott (2005). *Lexicalization and Language Change*. Cambridge: Cambridge University Press.
- Carvalho, Maria José (2006). *Documentação medieval do mosteiro de Santa Maria de Alcobaça (sécs. XIII-XVI). Edição e estudo linguístico*. Dissertação de doutoramento, Universidade de Coimbra.
- Heine, Bernd (1993). *Auxiliaries. Cognitive Forces and Grammaticalization*. New York/Oxford: Oxford University Press.
- _____, Bernd, Ulrike Claudi and Friederike Hünemeyer (1991). *Grammaticalization. A Conceptual Framework*. Chicago and London: The University of Chicago Press.
- _____, Bernd (2002). On the role of context in grammaticalization. In: Ilse Wischer and Gabriele Diewald (eds.), *New Reflexions on Grammaticalization*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 83-101.
- Hopper, Paul J. (1991). On some principles of grammaticization. In: Elizabeth Closs Traugott and Bernd Heine (eds.), *Approaches to Grammaticalization*, Vol I: *Focus on Theoretical and Methodological Issues*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 17-35.
- _____, Paul J. e Elizabeth Closs Traugott [2003² (1993)]. *Grammaticalization*. Cambridge: Cambridge University Press.
- Lehmann, Christian (1995). *Thoughts on Grammaticalization*. München – Newcastle: LINCOM Studies in Theoretical Linguistics, 1.
- _____, Christian (2002). New reflexions on grammaticalization and lexicalization. In: Ilse Wischer and Gabriele Diewald (eds.), *New Reflexions on Grammaticalization*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 1-18.
- Lima, José Pinto de (1997). Caminhos semântico-pragmáticos da gramaticalização: o caso de “embora”. In: A. M. Brito et al.(ed.). *Sentido que a Vida Faz. Estudos para Óscar Lopes*. Porto: Campo das Letras, pp. 646.

- Lorenzo, Ramón (1977). *La traducción gallega de la Crónica General y de la Crónica de Castilla*. Edición crítica anotada, con introducción, índice onomástico y glosario. Orense: Instituto de Estudios Orensanos “Padre Feijoo”, vol. II.
- Maia, Clarinda de Azevedo (1997). *História do galego-português. Estado linguístico da Galiza e do Noroeste de Portugal desde o século XIII ao século XVI (Com referência à situação do galego moderno)*. Reimpressão da edição de 1986. Lisboa: FCG e JNICT.
- Mettmann, Walter (ed.). (1972). *[Afonso X], Cantigas de Santa Maria*. Coimbra: Acta Universitatis Conimbrigensis, vol. IV (Glossário).
- Moreno Cabrera, Juan C. (1998). On the relationships between grammaticalization and lexicalization. In: Anna Giacalone Ramat e Paul J. Hopper (eds.), *The Limits of Grammaticalization*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 211-227.
- Nunes, José Joaquim (1989). *Compêndio de gramática histórica portuguesa. (Fonética e Morfologia)*. 9ª edição. Lisboa: Clássica Editora.
- Rodríguez, José Luís (1996). Sobre algumas formas e locuções conjuntivas nos textos galegos medievais. In: M^a do Carmo Henriques Salido (ed.), *Actas do IV Congresso Internacional da Língua Galego-portuguesa na Galiza. Homenagem a Ferdinand de Saussure* (Vigo, 28 Outubro-1 Novembro 1993). A Coruña: AGAL, pp. 189-201.
- Silva, Rosa Virginia Mattos e (1984). “Pero” e “porém”: Mudanças em curso na fase arcaica da língua portuguesa. In: *Boletim de Filologia*, vol. 29, pp. 129-152.
- Traugott, Elizabeth Closs (1995). Subjectification in Grammaticalisation. In: Dieter Stein e Susan Wright (eds.), *Subjectivity and Subjectivisation. Linguistic Perspectives*. Cambridge: Cambridge University Press, pp. 31-54.
- _____, Elizabeth Closs. (1999). The Rhetoric of Counter-Expectation in Semantic Change. In: Andreas Blank e Peter Koch (eds.), *Historical Semantics and Cognition*. Berlin, New York: Mouton de Gruyter, pp. 177-196.
- _____, Elizabeth Closs e Bernd Heine (eds.) (1991). *Approches to Grammaticalization*, vol I. Amsterdam/Philadelphia: John Benjamins Publishing Company.