

UN CORPUS PARA O ESTUDO DO GALEGO MEDIEVAL: O *TMILG (TESOURO MEDIEVAL INFORMATIZADO DA LINGUA GALEGA)*

Paulo Martínez Lema

Instituto da Lingua Galega. USC

1. O *TMILG* COMO PROXECTO: OBXECTIVOS, PLANTEL E PREVISIÓNS DE FUTURO

O *TMILG* (*Tesouro Medieval Informatizado da Lingua Galega*) constitúe hoxe por hoxe o único buscador on-line dispoñible para a lingua galega no seu estadio medieval, se ben debe enmarcarse dentro dun proxecto moito máis amplo e ambicioso, o *Tesouro Medieval Galego-Portugués (TMGP)*, que pretende pór a disposición do usuario a inmensa masa textual producida en Galiza e Portugal ao longo da Idade Media. Os criterios organizadores deste corpus son basicamente o glotolóxico e o xeográfico, e é a interacción dos mesmos a que nos permite establecer cinco subcórpora diferenciados: á marxe do *TMILG* propiamente dito, temos que falar do *TMILC-G* (textos galegos en lingua castelá, que nun futuro constituirán o *Corpus Histórico do Castelán de Galicia* ou *CHCG*), do *TMILL-G* (textos galegos en lingua latina), o *TMILP* (textos portugueses en lingua portuguesa) e finalmente o *TMILL-P* (textos portugueses en lingua latina). O propio procesamento de textos para o *TMILG*, moitos dos cales están total ou parcialmente redactados en latín (no caso dos máis temperáns) e en castelán (nos dos máis serodios), permite ir construíndo de xeito progresivo e simultáneo estes cörpera, dándolle sempre especial preeminencia ao *TMILP* por razóns lingüísticas e históricas evidentes. Neste sentido, están abertas senllas e frutíferas vías de colaboración cos equipos dos profesores portugueses Maria Francisca Xavier (Universidade Nova de Lisboa) e João Malaca Casteleiro (Academia das Ciéncias de Lisboa).

O *TMILG* naceu no seo do Instituto da Lingua Galega no ano 1992, se ben algúns dos membros do equipo inicial traballan hoxe noutras universidades galegas. O desenvolvemento deste proxecto implica labores diversos que van do arrequentamento e diversificación do corpus (in-

corporación de novas fontes) ao procesamento e revisión minuciosa dos textos de acordo cunhas pautas previamente establecidas e sometidas a unha constante actualización por parte do grupo. Tendo isto en conta, non debe estrañar que no *TMILG* participasen e participen persoas de extraccións e condicións diversas: bolseiros, contratados, investigadores da Universidade de Santiago de Compostela e mesmo voluntarios procedentes principalmente dos derradeiros cursos da titulación de Filoloxía Galega, así como dos distintos programas de doutoramento directa ou indirectamente relacionados coa lingua galega medieval.

O proxecto realízase desde as súas orixes baixo a dirección de Xavier Varela Barreiro, se ben na actualidade conta con dous bolseiros responsables de cadansúa sección: Ricardo Gutiérrez Pichel (fontes documentais e edición de textos) e Paulo Martínez Lema (codificación do corpus). Tamén figuran a día de hoxe como bolseiros Rocío Dourado Fernández, Lucía Doval, Iria Pérez Mariño e Iago Santiso Fernández, encargados principalmente da revisión e corrección das coleccións documentais asignadas en cada caso e en función das necesidades máis inmediatas do proxecto.

É xusto recoñecer que o *TMILG* non sería un proxecto posible nin viable sen o apoio de distintas institucións, entre as cales hai que salientar a propia Universidade de Santiago de Compostela, que xa desde as orixes cedeu un local de traballo axeitado no Instituto da Lingua Galega, así como o propio persoal investigador. A nivel estritamente económico cómpre destacar as achegas da Secretaría Xeral de Política Lingüística, que veu prestando apoio neste sentido tamén desde as fases máis embrionarias do *TMILG*.

O *TMILG* foi posto en rede por vez primeira no ano 2004 contando coa axuda técnica e informática do Centro Ramón Piñeiro (CIRP), que o tivo accesible no seu servidor até setembro do ano 2006. Ao longo dese período rexistráronse varios centos de usuarios regulares, concretándose máis de 23.000 consultas en pouco menos de 4.500 sesións, dato que deixa constancia do interese e utilidade (para investigadores principalmente, mais non só) dun produto coma o *TMILG*. En decembro de 2007 o buscador trasladouse á páxina web do ILG (<http://ilg.usc.es/tmilg/>), incorporando nesta nova versión novos modos de busca, de asociacións gráficas e de cálculo de estatísticas. Desde o mes de febreiro de 2008, esta nova versión do buscador vén funcionando de xeito ininterrompido e serviu xa para a realización dun elevado número de procuras, o cal dá boa conta da importancia que unha ferramenta deste tipo ten hoxe por hoxe para os investigadores.

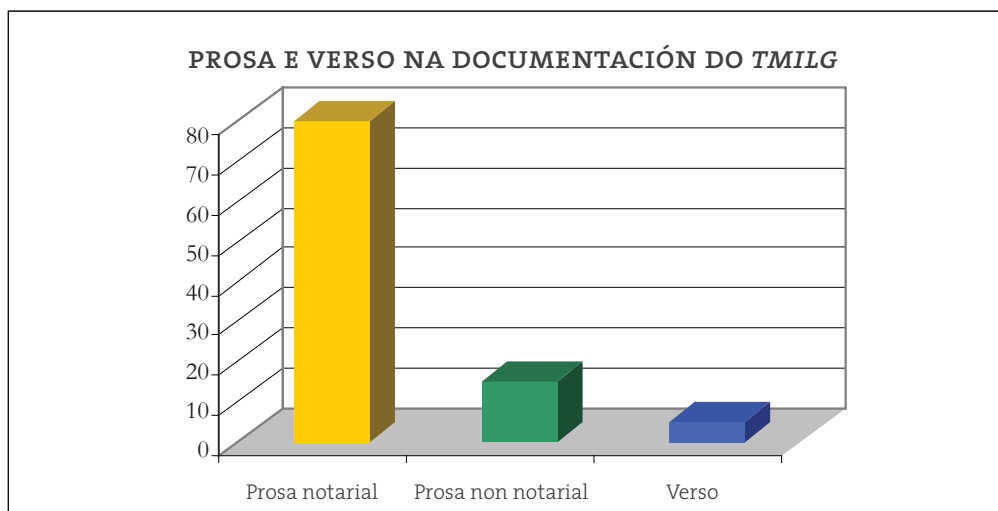
As expectativas de futuro do *TMILG* son certamente prometedoras, sobre todo no tocante á multiplicidade de proxectos que, directa ou in-

directamente, derivan del. No terreo estritamente lexicográfico botáronse a andar o *Diccionario Histórico da Lingua Galega* e mais o *Diccionario Medieval Galego-Portugués*, este último en colaboración con equipos da Universidade Nova de Lisboa. Tamén se está a traballar na *Dialectoloxía Medieval Galego-Portuguesa*, no *Inventario Toponímico da Galiza Medieval (ITGM)*, que se pode consultar xa no seguinte enderezo web: <http://ilg.usc.es/itgm/>, na *Gramática Histórica da Lingua Galega* e na edición paleográfica de distintas obras medievais xa incorporadas ao *TMILG*, mais que nalgúns casos reclaman unha lectura máis actualizada ou, cando menos, máis axeitada ás esixencias que presentan actualmente os estudos de lingüística histórica. Algunhas desas obras, en fase avanzada de transcripción, son o *Tratado de Albeitaria*, os *Miragres de Santiago* e as *Cantigas de Santa María*. Todo isto sen esquecer que o *TMILG*, xunto co *TILG*, os dous *Dicionarios de Dicionarios* (o medieval e mais o moderno) e o Seminario de Lingüística Informática da Universidade de Vigo, participará na construción dunha plataforma de recursos informáticos para a investigación da lingua galega (RILG), co patrocinio do Ministerio de Educación e Ciencia e da Consellaría de Innovación e Industria da Xunta de Galicia.

2. O CORPUS DO *TMILG*: ESTRUCTURA E CARACTERÍSTICAS

2.1. *Xéneros textuais*

O corpus do *TMILG* abrangue a día de hoxe uns 16.000 documentos, distribuídos á súa vez nun total de 82 obras e representativos das tres grandes categorías en que se articula a produción textual da Galiza medieval: a prosa notarial, a prosa non notarial e a poesía (verso). Mais, tal e como comprobamos nesta gráfica, o peso cuantitativo de cada unha delas é moi desigual:



Dos tres grupos que vimos de mencionar, o que maior volume de texto achega é sen dúbida a prosa notarial, que constitúe por si soa case un 80% do total. Esta abundancia cuantitativa ten o seu correlato nunha grande riqueza de subtipos textuais, se ben esa diversidade pode reducirse en última instancia a cinco grandes grupos (Vázquez Bertomeu 2001). Eses grupos son os seguintes:

(a) documentos referentes á adquisición e xestión do patrimonio: neste grupo, quizais o máis amplo e variado, incluímos vendas, traspasamentos, desembargos, aforamentos, arrendamentos, alugueiros, nominacións, almoedas, doazóns, concambias... Cómpre lembrar que moitos destes tipos presentan á súa vez subtipos específicos, como ocorre p.ex. nos aforamentos ou contratos forais, categoría moi ampla que pode xerarse en subconxuntos textuais máis reducidos como os traspasamentos de foro, o subaforamento ou a renovación de foro, entre outros;

(b) documentos relativos ao mundo mercantil e ás transaccións comerciais: gardas, peñoramentos, fianzas... formarían este segundo subconxunto textual;

(c) documentos administrativos: memoriais, inventarios de bens, apeos, relacións...

(d) documentos pertencentes ao ámbito familiar e doméstico: nesta categoría incluímos textos como as mandas, testamentos, codicilos, dotes...

(e) escritos de tipo xeral: testemuños, procurazóns, poderes, avinzas, sentenzas, renunciamentos e quitazóns, albalás, outorgamentos, demandas, frontas...

A estas cinco categorías xerais podemos engadir aínda outras dúas representadas por documentos de características temáticas bastante peculiares, mais que veñen especificamente singularizados polo contexto institucional en que se xeraron. Referímonos, por un lado, aos documentos reais (privilexios, acoutamentos, mandatos, ordenanzas...), e, polo outro, aos documentos eclesiásticos e especialmente aos monásticos (cartas de irmandade, decretos de anexión, repartimentos, restauracións monásticas...).

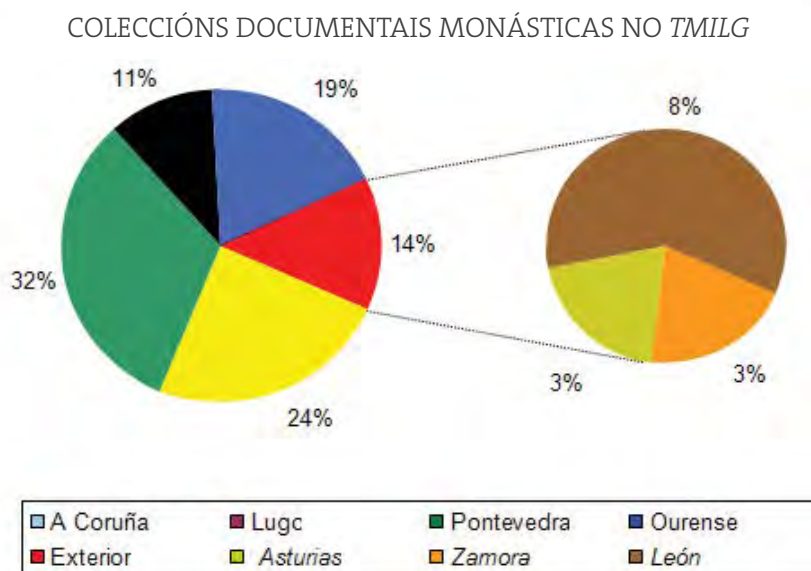
Desde o punto de vista dos focos xeradores de documentación, o *TMILG* permítelle ao usuario acceder á documentación elaborada tanto nos núcleos eclesiásticos (catedrais, mosteiros, parroquias, sínodos) como nos civís (concellos, universidades, documentos de carácter privado e outros textos de natureza diversa). Porén, é o primeiro grupo, o da prosa notarial relixiosa e máis concretamente a redactada nos mosteiros, a que goza dunha maior presenza numérica dentro do corpus. Neste sentido, están xa incorporadas ao *TMILG* un total de 32 coleccións documentais procedentes de diversos mosteiros, dos cales 12 radican na provincia de Lugo, 9 na da Coruña, 7 na de Ourense e 4 na de Pontevedra. Ademais, están xa revisadas, preparadas e listas para a súa inminente inclusión varias coleccións diplomáticas de mosteiros pertencentes aos territorios galegofonos de Asturias, Zamora e León. Estas coleccións suman un total de 5 obras, o que converte a Galiza exterior nunha área xeográfica e lingüística xa notablemente representada no *TMILG*. Porén, o constante traballo de enriquecemento das fontes bibliográficas e de procesamento dos textos extraídos delas fai que estas cifras sexan sempre provisionais e susceptibles de actualizacións periódicas¹.

Mosteiro	Concello	Provincia
San Antoniño de Toques	Toques	A Coruña
San Clodio do Ribeiro	Leiro	Ourense
San Martín de Castañeira	Puebla de Sanabria	Zamora
San Martiño de Vilourente	Mondoñedo	Lugo
San Miguel da Colleira	O Vicedo	Lugo
San Miguel de Bóveda	Amoeiro	Ourense

1. É necesario deixar clara esta circunstancia, pois, por poñer unicamente un exemplo, as cifras relativas á provincia da Coruña poderían engrosarse grazas aos textos procedentes de mosteiros como San Martiño de Cánduas (Ponteceso) ou San Cristovo de Dorneá (Boimorto), ou incluso con documentos serodios (finais do s. XV) que nos relatán procesos de reforma monástica que afectaron a cenobios como o de San Mamede de Seavia (Coristanco). Ao tratárense de obras aínda en proceso de revisión e conteren total ou maioritariamente textos en latín e/ou en castelán non foron tidas en conta neste cómputo.

Mosteiro	Emprazamento	Provincia
San Paio de Antealtares	Santiago de Compostela	A Coruña
San Pedro de Montes	Ponferrada	León
San Pedro de Ramirás	Ramirás	Ourense
San Pedro de Rocas	Esgos	Ourense
San Pedro de Soandres	A Laracha	A Coruña
San Pedro de Vilanova de Dozón	Dozón	Pontevedra
San Salvador de Bergondo	Bergondo	A Coruña
San Salvador de Celanova	Celanova	Ourense
San Salvador de Lourenzá	Lourenzá	Lugo
San Salvador de Pedroso	Narón	A Coruña
San Salvador de Sobrado de Trives	Trives	Ourense
San Salvador de Vilar de Donas	Palas de Rei	Lugo
San Salvador e San Nicolás de Cins	Oza dos Ríos	A Coruña
San Vicenzo de Pombeiro	Pantón	Lugo
San Xoán de Caaveiro	A Capela	A Coruña
San Xurxo de Codeseda	A Estrada	Pontevedra
Santa Clara de Santiago	Santiago de Compostela	A Coruña
Santa María de Belvís	Santiago de Compostela	A Coruña
Santa María de Carracedo	Cacabelos	León
Santa María de Castro de Rei de Lemos	Paradela	Lugo
Santa María de Ferreira de Pallares	Guntín	Lugo
Santa María de Ferreira de Pantón	Pantón	Lugo
Santa María de Oseira	San Cristovo de Cea	Ourense
Santa María de Valdeflores	Viveiro	Lugo
Santa María de Vilanova dos Oscos	Vilanova dos Oscos	Asturias
Santa María do Camiño	Pontevedra	Pontevedra
Santiago de Ermelo	Bueu	Pontevedra
Santiago de Mens	Malpica	A Coruña
Santo Domingo de Viveiro	Viveiro	Lugo
Santo Estevo de Chouzán	Carballedo	Lugo
Santo Estevo de Ribas de Sil	Nogueira de Ramuín	Ourense
Santos Xusto e Pastor de Toxos Outos	Lousame	A Coruña
Veiga de Espiñareda	Veiga de Espiñareda	León

No seguinte gráfico podemos observar máis detalladamente a relación porcentual existente entre as distintas procedencias xeográficas da documentación cenobítica do *TMILG*, tomando como referencia o marco administrativo actual. Deste xeito podemos observar algúns datos bastante congruentes tanto co decorrer da historia medieval galega como, máis especificamente, coas tendencias xeográficas dos asentamentos monásticos no Noroeste peninsular: por exemplo, comprobamos que a provincia que maior cantidade de coleccións documentais achega é Lugo, e máis en concreto a súa área meridional, que, xunto cos territorios máis setentrionais da provincia de Ourense artellados ao redor do río Miño, constitúe unha área especialmente rica no que a institucións monásticas se refire: a Ribeira Sacra. Por outra banda, resulta cando menos rechamante a precariedade documental que presenta a provincia de Pontevedra (4 coleccións), superada de feito polos territorios galegófonos hoxe administrativamente asturianos e castelán-leoneses (5 coleccións). A provincia da Coruña, pola súa banda, proporciónalle ao *TMILG* 9 coleccións documentais, se ben hai que ter en conta o considerable peso cuantitativo e cualitativo que teñen os mosteiros radicados en Santiago de Compostela, capital eclesiástica de extraordinaria relevancia neste sentido.



No que se refire ao outro grande bloque de prosa notarial, o constituído polos textos civís, domina claramente a documentación de tipo particular, sendo xa moito menos numerosa a procedente doutros ámbitos como o universitario e o municipal (soamente os concellos de Santiago, Noia, Pontevedra, Ourense e A Coruña achegan produción textual relevante ao *TMILG*).

O segundo grupo en importancia cuantitativa é o da prosa non notarial, que comprende aproximadamente o 15% do total dos textos. A prosa non notarial está representada por obras ou fragmentos de obras pertencentes a cinco grandes bloques temáticos ben coñecidos e delimitados:

(a) prosa literaria: pertencen ao xénero da narrativa extensa a *Crónica Troiana* (CT), a *Historia Troiana* (HT) e mais o fragmento conservado do *Livro de Tristan* (LT);

(b) prosa histórica: como sabemos, a historiografía medieval coñecida en lingua galega nútrese exclusivamente de traducións de orixinais procedentes doutros ámbitos lingüísticos, nomeadamente o castelán. Soamente dispomos de tres textos que reúnan as características definitivas deste xénero: a *Xeral Historia* (XH), a *Crónica Xeral* e *Crónica de Castela* (TC) e mais a *Crónica de Santa María de Iria* (CI);

(c) prosa relixiosa: no *TMILG* rexístranse tan só dúas obras deste tipo, os *Miragres de Santiago* (MS) e un fragmento do tratado *Sobre a simonía* (SaS). Descoñecemos, cando menos a día de hoxe, obras prosísticas galegas medievais pertencentes ao xénero da haxiografía, a catequística ou o sermón, por exemplo;

(d) prosa técnica: no noso corpus inclúense dúas obras que, polo seu carácter descritivo e directivo, poden clasificarse sen maiores problemas dentro desta categoría, se ben cada unha delas desenvolve temáticas moi afastadas entre si. Por unha banda, a *Arte de Trobar* (AT) inscribítese dentro do ámbito da teoría poética e, a un nivel máis amplo, no mundo das ciencias sociais e humanas; en cambio, o *Tratado de Albeitaria* (TA) céntrase no terreo da veterinaria e, xa que logo, pode considerarse un manual de ciencias aplicadas, o único existente a día de hoxe na nosa tradición prosística medieval;

(e) prosa xurídica: aínda carecendo de tratados xurídicos propiamente ditos, a prosa galega medieval si conta cun amplo repertorio de ordenamentos e códigos legais, todos eles xa dispoñibles no *TMILG*: referímonos aos fragmentos conservados do *Foro Xuzgo* (FX), *Flores de Dereito* (FD), *Foro Real* (FR), *Glosas de Textos Legais* (GTL) e o *Ordenamento de Alcalá de Henares* (OA), así como a versión completa conservada dos *Foros de Castelo Rodrigo* (FCR). A todas estas obras hai que engadir aínda os distintos fragmentos das sete *Partidas* de Afonso X, na súa tradución ao galego.

Finalmente, os textos en verso (todos eles pertencentes ao subxénero lírico nas súas distintas variantes), a pesar da súa innegable relevancia cualitativa dentro da produción escrita galega medieval, son os menos numerosos do noso corpus textual e supoñen aproximadamente un 5% do mesmo. Este conxunto pode dividirse á súa vez en dous grupos ben coñecidos e en cuxa caracterización interveñen o factor cronolóxico e a tradición estética na que se inscriben os textos:

(a) lírica trobadoresca, que abrangue *grosso modo* as derradeiras décadas do s. XII, a totalidade do s. XIII e a primeira metade do s. XIV. No *TMILG* podemos consultar os distintos subxéneros da vertente profana, maioritariamente agrupados baixo as siglas LP (*Lírica Profana*), con excepcións como o *descort* plurilingüe de Raimbaut de Vaqueiras (DPRV), xoia poética que inclúe unha estrofa en galego, ou o texto *Us Castellans* (UC) de Raimón Vidal de Besalú. Completan este inxente corpus poético as *Cantigas de Santa María* (CSM), tanto as narrativas como as propiamente líricas ou *de loor*, e mais a *Cantiga de Valcavado* (CV), as dúas únicas mostras poéticas de temática estritamente relixiosa con que conta a nosa tradición literaria medieval;

(b) lírica da decadencia (CGC), ou, atendendo á denominación máis comunmente aceptada e empregada, *escola galego-castelá*, que se desenvolve ao longo dos séculos XIV e XV.

Deste xeito, no *TMILG* encontramos perfectamente representadas as diversas parcelas textuais cultivadas na Galiza medieval, quer no terreo da prosa, quer no do verso. Con todo, e tal como pode observarse no cadro seguinte, (elaborado a partir do modelo de codificación do CORDE exposto en RAE 2001), non deixa de ser considerable o número de xéneros e subxéneros que non contaron co debido desenvolvemento e para os que, en consecuencia, carecemos de obras representativas. Mais esta situación está suxeita en todo caso ás alteracións que poderían derivar da descuberta e difusión de textos até o de agora inéditos.

PROSA		
Narrativa	Breve	
	Extensa → <i>Novela</i> (CT, HT, FT)	
Ciencia e técnica	Ciencias sociais e humanas → <i>Literatura</i> (AT)	
	Ciencias aplicadas → <i>Veterinaria</i> (TA)	
Sociedade		
Relixión	Textos bíblicos e glosas	
	Haxiografía	
	Mística e ascética	
	Catequística	
	Sermóns	
	Outros (MS, SS)	
Historia	Historiografía (TC, GE, CI)	
	Biografía	
Dereito	Ordenamentos e códigos legais (FX, FD, FR, GTL, OA, Partidas)	
	Tratados xurídicos	
	Outros → <i>Prosa notarial</i>	
Lírica		
Teatro		
Didáctica		
VERSO		
Narrativa	Tradicional	
	Culta	Épica
		Relixiosa (CSM-cantigas narrativas)
	Didáctica	
Lírica	Tradicional	
	Culta	Individual (CSM-cantigas de loor)
		Obras colectivas (LP, CGC, DPRV, UC)
		Outros
Dramática		

2.2. Cronoloxía e glotoloxía dos textos

Desde o punto de vista da cronoloxía, o corpus do *TMILG* cobre un período amplísimo como é o comprendido entre os séculos VIII (ano 700) e XVII (ano 1600). Neste sentido, o documento máis antigo co que contamos a día de hoxe no noso corpus é unha doazón procedente do mosteiro de Santa María de Sobrado dos Monxes, datada no ano 787 e editada por López Ferreiro no seu *Galicia Histórica* (capítulo 113) (López Ferreiro 1901:

626-627). Aínda que no *TMILG* podemos encontrar representadas diversas linguas, o certo é que soamente o galego / galego-portugués, o latín e o castelán teñen verdadeira presenza como linguas vehiculares preferentes dos textos. Deste xeito, por volta de 9.000 dos documentos consultables (un 56% do total) están escritos en idioma galego, fronte aos aproximadamente 4.500 (un 28%) redactados en latín e os escasos 2.500 (o 15% restante) para os que se escolleu a lingua latina.

3. MICROESTRUTURA DO TMILG: O PROCESAMENTO DOS TEXTOS

Unha das fases de maior relevancia do proceso de tratamento dos textos consiste na previa etiquetaxe dos mesmos, aplicándolles un encabezamento no que se fan constar distintos ítems informativos, como son:

(a) o rexesto: inclúese o rexesto literal redactado polo editor para cada documento, mais prescindindo de informacións complementarias como a cota arquivística ou os detalles codicolóxicos, por exemplo;

(b) o título da obra: como puidemos comprobar nas epígrafes precedentes, no *TMILG* combínase o título extenso da obra (regularizado e traducido sempre ao galego) cun título en clave formado xeralmente a partir das iniciais daquel (p.ex., *Historia Xeral* = HX);

(c) a información autorial, *i. e.*, o nome do autor de cada documento: neste punto faise constar, en primeiro lugar, o nome do autor con todos os seus posíbeis atributos tal e como figuran no texto (p. ex., *Johannes Petri, notarius concilii Noye juratus*), mentres que nun segundo momento rexistramos unicamente a cadea antroponímica illada, sen ningún tipo de información complementaria (*Johannes Petri*). Contémplase tamén a posibilidade da autoría múltiple coas súas diversas variantes, entre as cales salientan como máis habituais a presenza dun notario e dun escribán ao seu servizo (caso no que figura sempre en primeiro lugar o nome do notario), a participación de varios notarios ao mesmo nivel na xénese do documento, etc. Todas estas opcións contan cun sistema de codificación específico dentro da etiqueta *Autoría*;

(d) información cronolóxica: sinálanse o século e o ano ou rango de anos concreto no que foi redactado cada documento. Tamén fixemos constar o ano no que o documento en cuestión foi copiado ou trasladado, información que explica moitas das distorsións léxicas, gráficas ou doutro tipo observables nos textos;

(e) natureza do texto, i.e., se o documento en cuestión é un orixinal, unha copia, un traslado, un rexesto, unha tradución...

(f) información lingüística: indícase a lingua do texto mediante unha serie de claves numéricas específicas, e nos documentos traducidos exprésase ademais a lingua na que foi redactado o orixinal.

Na revisión do material textual ponse especial atención en illar, mediante os marcadores oportunamente establecidos, aqueles treitos, capítulos ou mesmo textos enteiros escritos en linguas distintas do galego. Deste xeito, o corpus do *TMILG* pode dividirse en varios subconxuntos desde o punto de vista glotolóxico, un dos cales é precisamente o dos documentos híbridos. O *TMILG*, como xa indicamos, abrangue desde textos moi temperáns, con fortísima presenza do latín, até textos moi serodios, nos que o castelán consegue hexemonizar xa a práctica escrita. Fica polo medio unha importante masa textual ao longo da cal imos asistindo á progresiva emerxencia e consolidación do romance galego-portugués como lingua escrita, sempre en tensión coa tradición escrituraria latina precedente e coa paulatina intromisión do castelán a partir do s. XIV. Xa que logo, a porcentaxe de textos híbridos resulta considerable, e o *TMILG* permítenos explotar as potencialidades que esta realidade lle ofrece ao investigador en eidos moi diversos que van desde a lexicografía á sociolingüística, pasando pola fraseoloxía ou a onomástica. Por exemplo, nun documento pertencente á colección da Casa de Altamira e datable nas primeiras décadas do s. XVI, Gutierre de Hevia, alcaide das fortalezas de Navia e Burón polo Conde de Altamira, quíxase do maltrato que recibe por parte do seu señor. A protesta fórmase nun perfecto castelán, mais nun determinado momento Gutierre de Hevia opta pola expresividade dun dito autóctono na lingua do país:

Por cierto no fuera mucho que onbre tan maltratado desesperara en algo de lo que arriba digo porque ay un enxemplo en Galizia que dize *el can con rabia en seo dono traba*.

Noutras ocasións, a irrupción do discurso oral nos documentos ilustra aspectos máis lúdicos da lingua da época. Nun requirimento feito pola abadesa de San Miguel de Bóveda en 1467 menciónase o xudeu Abraán de León, recadador dos Condes de Ribadavia, e dísenos que “quando se partira de ali que disera o dito judio: ¡como agora enganei a señora abadesa con un rabo de pescado!”;

(g) información xeográfica: nesta etiqueta indícase a demarcación ou demarcacións onde exercía o seu labor o notario responsable da creación do documento. Cando se descoñece tal dato, figura tan só o lugar con-

creto no que foi redactado o texto. Este apartado, como se comprenderá, resulta de especial interese para os historiadores e para os toponimistas, pois sempre se procura representar a forma lingüística exacta coa que aparece nomeada no documento esa referencia xeográfica concreta, así como, na medida do posible, contextualizala dentro dos parámetros administrativos que lle corresponderían na actualidade (provincia, concello, freguesía...). Por outra banda, cómpre lembrar que a vinculación dun notario cunha determinada xeografía ou xurisdición pode levar aparellado o emprego duns hábitos escriturais particulares a nivel léxico, gráfico e/ou morfosintáctico. Por tanto, o estudo desas relacións pode ser de grande utilidade para ir artellando un modelo de dialectoloxía medieval do romance galego-portugués;

(h) tipoloxía documental do texto: sinálase se o documento en cuestión responde ás características temáticas e/ou estruturais dun contrato de aforamento, dunha doazón, dunha venda... Co fin de ir creando un repertorio amplo e representativo faise constar sempre o nome que se lle dá no texto á acción xurídica da que este emana, mais tamén as construcións sintácticas (moitas veces fosilizadas e convertidas en verdadeiras fórmulas ou lexías complexas asociadas a unha tipoloxía textual concreta) utilizadas para expresar a natureza da dita acción. Por exemplo, a tipoloxía *doazón* aparece representada nesta etiqueta por substantivos como *doaçõn, dationem, donationem, scripturam testamenti et donationis...* e por frases verbais como *dono et concedo, concedo et offero, do et offero, do et dono et offero...* nas que se observan as distintas posibilidades combinatorias duns elementos léxicos máis ou menos constantes e recorrentes, e desde logo moi ligados semanticamente á acción que xera e xustifica o documento en cuestión. Neste sentido, tamén se observa unha notable riqueza nos testamentos e mandas testamentarias: substantivos transparentes como *testamentum, mandatione, manda, ordinatione* ou *codeçilo* (coas múltiples variantes gráficas existentes en cada caso) conviven no corpo do documento con fórmulas como *uolo ut omnia mea ordinata remaneant, ordino et faço meu testamento*, etc. En definitiva, este apartado do encabezamento ofrece unhas posibilidades nada desprezables no estudo do léxico medieval e da súa vinculación semántica máis ou menos sistemática a unhas determinadas accións xurídicas.

Por outra banda, no estado actual dos traballos do TMILG, a variedade documental novidosa en relación ás clasificacións canónicas de historiadores e paleógrafos é significativa, como tamén o é a nomenclatura empregada para identificar cada un deses tipos, unha nomenclatura que en galego é moitas veces debedora da tradición castelá. Deste xeito, o pro-

cesamento de documentación para o *TMILG* permítenos rescatar termos até o de agora ausentes dos manuais e dicionarios ao uso, cando menos cos matices semánticos cos que concorren na nosa documentación medieval. Por exemplo, adoita denominarse *deslinde* ao documento no que se establecen os lindes dunha propiedade, herdade, freguesía ou couto. Tamén recibe este nome, mediante un proceso metonímico doadamente explicable, a propia marca ou sinal divisorio utilizado no deslindamento (*GDXL*, s.v. *deslinde*). Porén, nos textos medievais comparece un fermoso sinónimo de *deslindar*, como é *devisar* (coas súas múltiples variantes gráficas). Ramón Lorenzo (1968, s.v. *devisar*) rexistra xa este verbo, nas súas formas participiais, en documentos da segunda metade do s. XIII: “assj como estan marcadas et deuisadas... assi como jaz deuisada” (1259), “per terminos deuissados” (1274). Na nosa lexicografía moderna, Eladio Rodríguez atesta as formas *devisar* e *divisar* co significado ‘Divisar, señalar, referir, dividir’, e non debemos esquecer que o portugués moderno conserva a forma *devisar* co significado ‘Ordenar, planejar, construír de acordo com um plano’ (Houaiss 2001, s.v. *devisar*). Porén, nin o galego nin o portugués pareceron conservar o derivado sufixal *devisaçon*, aínda recollido por Viterbo (1798-1799) coa variante *divisaçom* (‘separação, divisa, termo, limite, partilha’) nun documento bracarense de 1501 e relativamente abundosa nos nosos textos antigos: así, na colección do mosteiro de San Salvador de Celanova encontramos o termo *devisaçiones* xa nun documento en castelán con data de 1293 (“los marcos et las devisaçiones”)², e nun texto algo posterior, concretamente de 1332, fálásenos da delimitación duns lugares “per marcos et devisações çertas”.

Aínda habería que engadir que o proceso de revisión de textos nos permite tamén identificar certos subtipos textuais caracterizados por matices xurídicos, estruturais e/ou mesmo lingüísticos que os diferencian do tipo textual máis xenérico no que se incardinan. Por exemplo, é ben coñecido o tipo textual *foro/aforamento*, mais na documentación do noso corpus poden recoñecerse subcategorías moi concretas e con entidade ben definida, como a *cesión de foro*, a *confirmación de foro*, a *demisión de foro*, a *devolución de foro*, o *traspasamento de foro*, o *subaforo*, o *nomeamento de foreiros*, etc. Nun documento procedente tamén do mosteiro de San Salvador de Celanova datado en 1426, un tal “Lourenço Miguelles, clerigo capellan do moesteiro de San Salvador d’Arnoya”, traspasa unhas herdades que el mesmo tiña aforadas polo devandito mosteiro, e exprésao nos

2. Curiosamente, o *CORDE* non dá ningunha ocorrencia do derivado *devisaçon* / *divisaçion* para o castelán, nin na época medieval nin na moderna, o que parece reafirmar máis aínda a singularidade desta forma.

seguintes termos: “parto et demito et traspaso”. Estas escollas léxicas entroncan o traspasamento de foro con outros tipos documentais profusamente representados no *TMILG* como son a *demitizón* ou a *renuncia*;

(i) información editorial: neste derradeiro apartado do encabezamento rexístrase a páxina da obra na que comeza o documento, así como o rango de páxinas que abrangue. Asemade, faise constar tamén o capítulo da edición co que se corresponde cada documento.

Este encabezamento veu funcionando, con certos reaxustes puntuais, ao longo destes case 20 anos de vida do *TMILG*, mais na actualidade o grupo de investigación está a desenvolver un novo e máis ambicioso etiquetario (DTD) mediante o que se recolla toda a información pertinente non só de cada documento en particular, senón tamén do conxunto da obra ou volume no que se recolle. Os ítems informativos contemplados nese novo encabezamento son esencialmente os mesmos aos que se veu recorrendo até o de agora, pero moito máis matizados e detallados. Ademais, incorporárase toda a información de tipo arquivístico e codicolóxico existente para cada documento e moitas veces accesible xa a través das propias edicións dos textos.

Xa dentro do documento en si, e nunha segunda e máis fonda fase de revisión, márcanse tamén, mediante os indicadores previamente convindos, todos e cada un dos elementos onomásticos: topónimos, antropónimos, títulos de obras, nomes de días e de meses, nomes de institucións... Esta práctica é a que permite desenvolver proxectos como o xa mencionado *Inventario Toponímico da Galiza Medieval*, que esixe un laborioso proceso previo de illamento, clasificación e identificación dos topónimos de cada documento que contribuíu a asentar unha metodoloxía de traballo xa ensaiada con éxito nalgúns traballos de doutoramento desenvolvidos por bolseiros do *TMILG*.

Como exemplo deste proceso de tratamento e etiquetaxe dos textos para o *TMILG* presentamos un documento pontevedrés do s. XV, pertencente a unha tipoloxía documental insólita até o de agora, o *contrabto de perdon*. O contido do escrito é certamente pintoresco, como tamén o son os termos e o léxico empregados: Pedro de Montes, natural do mosteiro de Aciveiro, comprométese a perdoar “todo mal, dano, pecado et copulaçon et adulterio” que a súa muller, Tareixa de Montes, cometera e puidera cometer até o 1 de abril seguinte, data na que ambos os dous deberían reencontrarse para “faser maridança segundo que manda Deus et Santa Iglesia”. Xa que logo, o interese do documento desborda os aspectos meramente lingüísticos para converterse nunha peza de indubidable valor histórico e sociolóxico.

[Rexesto do documento] [Pero de Montes, natural do mosteiro de Santa Maria de Acibeiro, perdoa á súa muller Teresa de Montes por ter abandonado o domicilio familiar e cometido adulterio, a condición que volte con el antes do próximo mes de abril]

[Acrónimo do título] [MNP]

[Título estandarizado] [Minutario Notarial de Pontevedra³]

[Volume] [-]

[Nome do autor (completo)] [Anónimo]

[Nome do autor (abreviado)] [Anónimo]

[Século] [15]

[Ano de redacción do documento] [1434]

[Ano de redacción do orixinal] [-]

[Mes de redacción do documento] [11 (= *novembro*)]>

[Orixinalidade] [O]

[Lingua do documento orixinal] [-]

[Lingua do documento] [1 (= *galego*)]

[Referencia xeográfica] [Pontevedra]

[Xeración trobadoresca] [-]

[Poesía / prosa (1 = prosa)] [1 (= *prosa*)]

[Xénero] [15 (= *prosa notarial*)]

[Tipoloxía documental] [Perdón (“contrabto de perdón”)]

[Páxina de inicio] [150]

[Rango de páxinas] [150-151]

[Capítulo ou número de documento (na edición de referencia)] [110]

O primeiro día do mes de *novembro*. Sabean todos que eu *Pero de Montes*, natural do mosteiro de *Santa María d'Asiveiro* que soo presente, quero et outorgo que por quanto vos *Tareixa de Montes* mña moller, filla de *Juan Branco* et de *Moor Gomes* sua moller, que sodes absente, vos absentastes de mña casa et poder sen meu mandado et liçencia a outra parte a donde vos quisestes, por lo qual vos non queredes tornar a mia casa et poder, nen eso mesmo non ousades de andar seguramente por la terra donde sodes natural, nen parescer ante mñ nen lugar donde eu de vos aja notiçia et parte et vos posa costrenjer por rigor de dereito, reçeando vos que por lo dito absentamento que asy fezeistes que eu vos queira demandar et acusar por rason que avedes cometido pecado de adulterio; por ende eu o dito *Pero de Montes* que asi soo presente, non costregudo por forza nen por engano rescebido, mas de mña livre et propia voontade et por lo amor de *Deus*, perdoe a vos a dita *de Montes* mña moller, que sodes absente asi como se fosedes presente, todo erro et mal et dano et interese que eu rescebese et ouvese por rason do dito absentamento que asi fezeistes, et eso mesmo todo pecado de adulterio que vos fezesedes et cometesedes et fezeistes et cometestes des lo tenpo que sodes mña moller et comigo casastes ata este presente día et que fezerdes et cometerdes, o que *Deus* non queira, desde oje que esta carta he feita ata o primeiro día do mes de *abril* primeiro que ven, que seera en este ano primeiro que verra de mill et quatroçentos et trinta et cinco anos, et qualquer copulaçon que fezesedes ou fezeistes ou festerdes ata o dito

3. Rodríguez González / Armas Castro (1992).

dia con qualquer persona en meu perjuiso. Et para esto así seer mas firme, vos a dita *Tareja de Montes* ata o dito primeiro día de abril avedes de viñr aa iglesia de *San Migell de Carvalleda*, et desde que ende fordes fasermo a saber para que vos et eu falemos en hüu seguramente et conçertandonos para <P 151> que en hüun façamos nosa vivenda, asi como por *Deus* et por dereito somos obligados de nos faser maridança segundo que manda *Deus* et *Santa Iglesia*, et non nos et eu como dito he de boo vos perdo et dou por libre et quita, para qui et por ante *Deus*, de todo mal, dano, pecado et copulaçon et adulterio que vos cometestes et fezeistes et cometerdes et festerdes en meu perjuiso ata o dito primeiro día de *abril*, et prometo et outorgo aa boa fe de *Deus* et sen mao engãno et por juramento que ende faço en este sinal de crus en lugar de *santos avangeos*, que eu que vos lo teña et conpra et agarde asi, et que en publico nen en scondido vos non beña nen pase, por min nen por outro, contra elo nen contra parte delo, et demais concertandonos ãna dita vivenda que, sen embargo do sobre dito, vos faça maridança do corpo et do aver ben et lealmente asi como marido a moller et por la via et forma que manda *Deus* et *Santa Iglesia*, et non no teendo et conprindo en todo asi et yndo et pasando contra elo, que eu seja enfames et fe perjuro, et me costamio de non ser de tal pecado asolto salvo por *noso señor o papa*, et demais que eu de et page por pena en nome de interese a vos a dita *Maria de Montes* dosentos froliins de ouro do cuño de *Aragon*, et apena pagada ou non que senpre este contrabto este firme, sobre lo qual renunçio as leis do dereito, dolo, engano, medo, força, endusimento et a todas las outras et dereitos canonicos et cevils et impereas et reas de que me eu ou outro por min pudeiramos ajudar para contra esto que dito he viñr, que quero et outorgo que me non vallan nen sejamos sobrelo nen sobre parte dlelo oydos nen resçebidos en juiso nen fora del; et por que esto seja çerto outorgey ende este contrabto de perdon ante o notario et testemoyas de juso scriptas. Que foi outorgado na villa de *Pontevedra*, ano, días, mes sobre ditos. Testemoyas que foron presentes, *Afonso Soares* caldeirerio morador ãna dita villa et *Garçia Corto* et *Gomes Ares* moradores en *Ribadavia* et *Gil Martines* clérigo de *San Justo de Avion* et outros. <P 152>

4. ALGUNHAS CONCLUSIÓNS

O *TMILG*, ao longo dos seus case catro anos en rede, converteuse nunha ferramenta indispensable para achegarse ás múltiples vertentes da cultura galega medieval, con especial relevancia das dimensións lingüística e histórica. Isto é así grazas ao abondoso material textual que constitúe o corpus base, mais tamén á súa diversidade xenérica e, en certo xeito, lingüística, o cal posibilita unha achega multidisciplinar a un período da nosa historia especialmente rico e aínda non totalmente explotado en todas as súas infinitas posibilidades.