
Gramática e teorização linguística em Portugal: a Gramática Filosófica de Jerónimo Soares Barbosa

Manuel Amor Couto
Universidade de Santiago de Compostela

Resumo:

A *Gramática filosófica da língua portuguesa* (1822) de Jerónimo Soares Barbosa é um dos textos fundamentais da tradição gramatical portuguesa. Em todo o seu texto e particularmente na *Introdução* podem-se encontrar observações de interesse sobre temas centrais da teorização linguística da época: a necessidade de uma gramática científica elaborada a partir de princípios rigorosos, as conexões das regras da linguagem com as regras lógicas, a divisão e conteúdo das partes da gramática e a urgência da reforma do ensino gramatical.

Palabras chave:

História da linguística portuguesa. Gramática portuguesa. Gramáticas filosóficas. Iluminismo português.

Abstract:

Jerónimo Soares Barbosa's A Philosophical Grammar of Portuguese is a key work in the studies of Portuguese grammar. In this text, and most particularly in the Introduction, we can find interesting observations about the linguistic theory of that time: the need for a scientific grammar based upon objective principles; the need to reform the teaching of grammar; the connection between the rules of language and the rules of logic, and the content and division of all parts of grammar.

Key words:

History of Portuguese Linguistics. Portuguese Grammar. Philosophical Grammar. Portuguese Illuminism.

Introdução

É o propósito deste artigo fazer uma exposição das ideias linguísticas presentes num dos gramáticos portugueses mais interessantes do iluminismo português, Jerónimo Soares Barbosa, autor de uma gramática filosófica que combina uma descrição linguística completíssima e singular com uma grande riqueza teórica e especulativa. Estas ideias estão especialmente presentes na sua *Introdução*, o

nosso principal objecto de análise neste artigo¹. Injustamente pouco conhecida e divulgada, a gramática de Soares Barbosa deixa-nos ver como a tradição gramatical portuguesa soube nesse momento histórico adaptar-se e aplicar os esquemas logicistas pensados primeiro noutras línguas. Por outro lado, a importância da conexão entre estruturas linguísticas e estruturas lógicas outorga a esta gramática um especial atractivo à hora de examinar a questão das relações entre língua e pensamento.

1. Jerónimo Soares Barbosa

Jerónimo Soares Barbosa era natural de Ansião, onde nasceu em 1737. Estuda no seminário de Coimbra, e em 1762 é ordenado presbítero. Em 1768 recebe o grau de Bacharel em Direito Canónico na Universidade de Coimbra. A sua vida está ligada ao ensino: de 1766 a 1790 exerceu a docência de retórica e poética no Colégio das Artes. Em 1789 foi nomeado sócio da Academia Real das Ciências de Lisboa e em 1792 designado inspector de ensino primário na circunscrição escolar de Coimbra. Um ano depois encarrega-se de promover e dirigir as edições de autores clássicos para uso das escolas. Em 1799 é nomeado vogal da Junta da Directoria Geral dos Estudos na Universidade de Coimbra.

Este interesse pela pedagogia concretava-se na renovação dos métodos de ensino da altura. Como muitos ilustrados, Soares Barbosa defende a aprendizagem da gramática latina desde a língua materna. A teoria gramatical devia também partir não da aplicação mecânica dos esquemas gramaticais latinos, mas da teoria da gramática geral ou filosófica, comum a todas as línguas. Os seus contributos mais importantes neste terreno foram a *Escola popular das primeiras letras* (1796), gramática elementar do português destinada à alfabetização nesta língua, acompanhada cada lição de instruções para os mestres, vem depois a sua obra fundamental, a *Grammática philosophica da lingua portugueza ou Principios da grammatica Geral applicados á nossa lingua-gem*, cuja introdução está datada em Coimbra a 24 de Junho de 1803, mas que não foi publicada pela Academia até 1822 e *As duas línguas* (1807), uma gramática filosófica comparada do latim e do português. Inclui nela apenas a “Etimologia” e a “Sintaxe”, as partes “lógicas” da língua. Soares Barbosa foi também tradutor de Quintiliano e Horácio para português (*Institutiones Oratoriae e Poética*)².

¹ Utilizamos a quarta edição, publicada em Lisboa em 1866. Embora sejam todas as edições do século XIX, a gramática de Soares Barbosa é manifestamente setecentista.

² Os dados sobre a sua vida e obra em Casteleiro (1980/81: 101-103).

2. O contexto da Gramática de Soares Barbosa: Ensino gramatical e reformas pombalinas

Embora redigida na fronteira entre os séculos XVIII e XIX e editada postumamente em 1822, a obra de Jerónimo Soares Barbosa responde em todos os aspectos ao substrato intelectual que caracterizou o iluminismo português, constituindo a sua máxima expressão no campo linguístico. Portugal não se viu à margem no século XVIII da curiosidade pelo progresso científico e tecnológico e as polémicas intelectuais que marcaram a Ilustração europeia, como as empreendidas contra a filosofia aristotélica, a crítica da doutrina do silogismo, a defesa da razão natural feita por Verney ou os debates entre inatistas de adscrição cartesiana e empiristas ou sensualistas da escola de Locke.

O século XVIII português aparece marcado pela importantíssima figura do marquês de Pombal, Sebastião José de Carvalho e Melo, um dos máximos representantes do absolutismo esclarecido. O impacto das reformas pombalinas deixa-se sentir em todos os âmbitos: administração, política colonial, produção industrial e agrícola e também orientação cultural e educativa. É neste último aspecto que as suas reformas atingiram a língua portuguesa.

Contudo, e antes de passar a enumerar as reformas mais importantes nesse apartado, importa sublinhar que não se trata de uma figura isolada, por mais importante que possa ser, mas do representante mais notável de uma série de reformadores que, no Portugal do século XVIII, questionaram muitos dos princípios ideológicos e culturais em que assentava a sociedade da época. Lembremos que a figura mais singular dos reformadores portugueses da época, Luís António de Verney, propugnava também nesta mesma altura uma reforma revolucionária dos métodos de ensino no seu *Verdadeiro método de estudar* (1749). A pesar de Soares Barbosa não reconhecer a sua dívida intelectual com Verney, não pode haver qualquer dúvida do vínculo entre o Barbadinho e a obra que analisamos, nomeadamente na questão da reforma ortográfica. Por outro lado, este pensamento crítico, autêntico espírito da época, não era mais do que o reflexo à portuguesa das grandes correntes intelectuais que por toda a Europa abalavam com as ideias tradicionais e que, em diferentes campos do saber e por diferentes autores (Rousseau, Montesquieu, Kant, Locke) punham os alicerces de um conhecimento e uma sociedade novas.

A preocupação pedagógica do Iluminismo português e europeu retoma as ideias de reforma pedagógica já existentes na Europa do século XVI (Vives, Melancton, Petrus Ramus...) e manifesta-se na época pombalina numa crítica radical do ensino jesuítico e dos seus procedimentos, baseados na filosofia aristotélica, no comentário textual, na *disputatio* entendida como simples exercício mecânico de habilidade lógico-formal e numa casuística arbitrária e probabilística. Em oposição a este ensino jesuítico estava o propugnado pelos iluministas, em que a teoria e a prática

pedagógica tinha como elemento essencial a ideia de método, uma das quatro divisões da Lógica (as outras três eram a origem e natureza das ideias, o juízo e o raciocínio). O método seguia o modelo das Matemáticas e da Geometria, ciências apoiadas num número reduzido de regras e exemplos de exactidão, ordem, coerência e claridade. O método era entendido como disposição dos argumentos, como instrumento que punha os recursos dialécticos e retóricos ao serviço da razão e permitia assim construir e comunicar o conhecimento. O ponto de partida era uma série de princípios definidos aprioristicamente que depois se demonstravam³. O método sintético (chamado também segundo as autores «de composição», «de doutrina» ou «natural») era considerado o mais ajeitado para a transmissão do conhecimento. Mas o método foi para os autores do século XVIII um instrumento, uma simples técnica, o que permitia a sua utilização por autores da escola empirista ou na terminologia da época «sensualista», que doutrinariamente estavam longe dos criadores do método, os racionalistas cartesianos do XVII.

Foi nesta época, depois de expulsos os jesuítas (1739), banida a sua Ordem e proibidas as gramáticas latinas jesuíticas que o poder político de diferentes países europeus promoveu reformas pedagógicas, introduzindo na Universidade e nos diferentes níveis do ensino estudos e metodologias modernas, que visavam consagrar oficialmente no campo concreto do ensino gramatical determinados manuais que se convertem em obrigatórios, como aconteceu em Portugal com a *Arte da Grammatica da Lingua Portuguesa* de Reis Lobato em 1770 ou em Espanha com a Gramática da Real Academia Espanhola de 1771⁴.

O próprio Soares Barbosa avaliava positivamente este dirigismo gramatical exercido desde o poder quando aludia ao alvará que, o 30 de Setembro de 1770, consagra como oficial a gramática de Reis Lobato (Barbosa 1866:XV). Pode-se ler íntegro este interessante documento na introdução de Carlos Assunção à sua edição da *Arte da Grammatica da Lingua Portuguesa*. As ideias aqui expostas têm idênticas preocupações e a mesma filosofia linguística que podemos ver na *Introdução* da gramática filosófica de Soares Barbosa :

...sendo a correção das línguas nacionaes hum dos objectos mais attendiveis para a cultura dos povos civilizados/.../ sabendo-a [a língua] por principios, e não por mero insticto, e habito,/.../ Sou servido ordenar, que os Mestres da lingua Latina, quando receberem nas suas Classes os discípulos para lha ensinarem, os instruaõ previamente por tempo de seis mezes, se tantos forem necessarios para a instrucção dos Alumnos, na Grammatica Portuguesa, com-

³ Calafate (2001, III: 212).

⁴ Sobre a gramática portuguesa de Reis Lobato veja-se Assunção 1997 e 2000. Para a Gramática da Academia Espanhola veja-se Sarmiento 1984: 9-28.

posta por Antonio José dos Reis Lobato, e por Mim aprovada para o uso das ditas classes, pelo methodo, clareza, e boa ordem, com que he feita (Assunção 2000: 27-28).

Vemos, portanto, uma convergência de interesses entre os propósitos centralizados, regulamentadores e antijesuíticos do poder da época e as sugestões de renovação pedagógica e gramatical formuladas por autores como Verney e Soares Barbosa.

3. Pedagogia e Gramática

A preocupação pedagógica em Soares Barbosa é perceptível em vários níveis:

- a) De um lado, como crítica negativa dirigida contra os péssimos métodos utilizados no ensino, quer fossem estes os métodos de transmissão do conhecimento usados pelos professores nas escolas, quer fossem os manuais escolares e a disposição com que estes apresentavam o objecto de estudo gramatical.

Os iluministas portugueses eram conscientes de que a metodologia para o ensino das línguas latina e vulgar precisava de reformas em profundidade. É por isso que se tentava substituir a bicentenária gramática latina de Manuel Alvares (1572), método clássico do ensino jesuítico, por outros manuais mais em conformidade com os novos tempos e ideias, como os de Verney⁵.

A crítica aos métodos de leitura e escritura era tema que atingia directamente a Soares Barbosa por ser autor de um tratado destinado a essa aprendizagem e ter desempenhado as funções de inspector de ensino primário. No texto da sua gramática fala de passagem de qual seria a prática pedagógica correcta, além de criticar o uso de documentos em letra manuscrita:

ensaiando-se debaixo da direcção de bons mestres a pronunciar-os depois, já soletrando-os, já syllabando-os, já pronunciando-os juntamente nos vocabulos e no discurso, e ligando tudo por meio de uma leitura certa, desembaraçada e elegante, o que nunca se conseguirá com os métodos e cartas informes, de que até agora se tem usado; mas sim com abecedarios e syllabarios exactos e completos, e principiando a leitura por cartas e livros de letra impressa, mais regular, mais uniforme, mais certa, e por isso mesmo tambem mais facil e mais propria para dar o leite das primeiras letras á tenra idade (Barbosa (1866: 37).

⁵ Ou, já no século XIX e um ano depois da primeira edição da Gramática de Soares Barbosa, uma gramática filosófica do latim, *Gramática Filosófica da Língua Latina, reduzida a Compêndio ou Método Suave de Ensinar a Aprender o Latim*, do mestre Diogo de Meneses (Calafate 2001, III: 241).

A crítica que Soares Barbosa faz das gramáticas (fundamentalmente as destinadas à aprendizagem do latim) é muito semelhante nos seus termos à crítica que Verney fizera na Carta Primeira do *Verdadeiro método de estudar*. Continua o nosso autor:

Porem deixada [a gramática] pelos philosophos nas mãos de homens, ou ignorantes ou pouco habeis, se reduziu a um systema informe e minucioso de exemplos e regras, fundadas mais sobre analogias apparentes que sobre a razão, a quem so pertence inquirir e assignar as verdadeiras causas da linguaagem, e segundo ellas ordenar a grammatica de qualquer lingua particular. D'aqui nasceram todas estas artes enfadonhas de grammatica latina, cheias de mil erros e de tantas excepções quantas são as regras. O que tudo repetido e copiado cegamente de idade em idade, sem nunca ter sido submettido a exame, sem o mesmo tambem foi servilmente applicado ás grammaticas das linguas vulgares. (Barbosa 1866: XII-XIII).

Esta crítica da casuística gramatical ia dirigida nomeadamente aos estudos de Prosódia e de Ortografia, os mais deturpados por uma prática escolar deficiente.

Uma das razões que Soares Barbosa alega para os gramáticos se ocuparem destes dois apartados, o prosódico e o ortográfico (e que implicava que, num sentido restrito, Ortoepia e Ortografia, por serem aspectos «mecânicos» e não lógicos, poderiam ficar fora do núcleo central da gramática, constituído pela Etimologia (Morfologia) e a Sintaxe) era a consciência da degradação que a prática escolar causara:

[...] de muito tempo a esta parte se tem entregado o ensino d'estas duas partes da grammatica Portugueza aos mestres de escola, pela maior parte pouco habeis. Porém d'aqui tem procedido os maus métodos, com que a primeira idade perde nas escolas boa parte do seu tempo, e gasta outra em aprender coisas que depois tem de desaprender ou reformar. É justo pois que a coisa torne a seu dono, e que os grammaticos tomem outra vez a si esta parte da grammatica que ensina a theoria dos sons, e tudo o que pertence á boa pronunciação e leitura da lingua, pois que tem sido tão mal desempenhada em mãos estranhas (Barbosa 1866: IX)

e mais adiante:

as regras mesmas da boa pronunciação e escriptura devem entrar no ensino da grammatica, para emendar muitos vicios que os mestres de primeiras letras, pela maior parte idiotas, não são capazes de corrigir (Barbosa 1866: XVI).

- b) De outro lado, como crítica de orientação positiva, mediante a sugestão de métodos descritivos e explicativos alternativos que, por responderem a modelos de clareza, simplicidade e elegância e terem carácter prático, fizessem mais fácil e rápido o processo de aprendizagem.

A aprendizagem linguística não tem a penas carácter instrumental. Conhecer a própria língua é, nas palavras do nosso autor, aprender a bem pensar⁶. Mas a aprendizagem da língua vulgar considera-se também como valioso instrumento auxiliar para o posterior domínio das linguas clássicas: “Esta theoria, applicada primeiro á propia lingua, percebe-se e comprehende-se muito mais facilmente do que applicada a linguas desconhecidas”. (Barbosa 1866: XV). Na ideia da precedência da língua materna sobre o latim parte-se do princípio pedagógico da construção do conhecimento desde o já sabido. Estas opiniões estão muito próximas das de Verney, que explica a sem-razão de ensinar gramática latina em latim⁷ ou, por falar no contexto galego, das reflexões do padre Sarmiento sobre o papel do galego no ensino⁸.

Por outro lado, o uso de um autêntico método científico no campo gramatical segundo os esquemas matemáticos e geométricos permitirá formular um conjunto de regras de número reduzido, baseadas umas nas outras e que podem ser decoradas facilmente⁹. As gramáticas modelo propostas afastam-se pois de maneira notável das anteriores.

4. Gramática: Ciência, método e lógica.

Soares Barbosa diferencia na sua *Introdução* entre o uso da faculdade da fala, própria de todos os seres humanos, e a ciência da linguagem, competência privativa dos gramáticos. A gramática considera-se arte, e é gramático quem ordena, classifica, relaciona e explica com os instrumentos proporcionados pelas leis da lógica e

⁶ “mostrar os principios gerais de toda a linguagem [...] formar assim uma lógica prática, que ao mesmo tempo que ensina a fallar bem a propria lingua, ensine a bem discorrer” (Barbosa (1866: XIV).

⁷ “É coisa digna de admiração que muitos homens deste Reino queiram aprender Francês, Tudesco, Italiano, de uma sorte, e o latim de outra muito diferente. Aprendem aquelas linguas com um Mestre que as fala ambas, e explica a língua incógnita por meio daquela que eles conhecem e falam; e com uma só Gramática se põem em estado de entenderem os autores bem, e, junto com o exercício, de falam Francês correntemente. E tomara que me dissesem por que não se deve praticar o mesmo no latim, e por que razão se haja de carregar a memória dos pobres estudantes com uma infinidade de versos latinos, e outras coisas que não servem para nada neste mundo” (Verney (1949: I, 141).

⁸ Veja-se por exemplo Monteagudo (1999: 266 - 284) ou Mariño (1998: 242-250).

⁹ “As regras propostas por este methodo reduzem-se a menos, porque se unem ao mesmo principio: percebem-se melhor, porque se sabe a razão d’ellas; e fixam-se mais na memoria, porque se ligam umas com as outras” (Barbosa (1866: XIV).

da razão uma linguagem que, deixada em estado natural, ficaria como simples uso espontâneo¹⁰. Neste sentido o gramático cria a língua no seu processo de ordenação, classificação e hierarquização da matéria prima linguística.

Tomar a lógica e o raciocínio como ponto de partida não significa, porém, que o gramático aplique de maneira mecânica e apriorística os princípios lógicos ao material linguístico, pois do que se trata é de estabelecer as leis da língua a partir de uma série de observações factuais¹¹ que permitam formular as regras particulares desse idioma, regras particulares que devem estar também em conformidade com as regras gerais que regem o pensamento.

A ideia que Soares Barbosa se faz tanto da historicidade das línguas como das diferentes etapas do estudo gramatical deriva-se da sua ideia evolutiva da língua: as línguas passaram por quatro estados sucessivos, cada um deles ligado a um sistema específico de representação escrita. Um primeiro estado foi o da representação pictográfica, considerada pelo nosso autor como método sumamente imperfeito e primitivo. O segundo foi a representação hieroglífica dos egípcios, originado numa estilização do sistema pictográfico em que debuxos com valor simbólico serviram como significante de uma ou várias coisas. O terceiro sistema de representação foi o ideográfico, conhecido na Europa do tempo através das informações que se tinham da cultura chinesa. A imagem natural, a conexão material que existia com o referente nos dois sistemas anteriores desaparece: o significante, já completamente estilizado, converte-se em simples símbolo. Sistema considerado como pouco económico, porém, pela necessidade de habilitar para cada ideia um símbolo, quando o número potencial de conceitos exprimíveis é infinito. O quarto estado é julgado como o mais perfeito e evoluído e corresponde com o sistema alfabético. Esta escrita, chamada literal, ocupou-se inicialmente da parte mecânica da linguagem, isto é, da série de sons que conformam cada língua e do sistema gráfico adoptado para a representação destes sons.

Esse carácter progressivo dos sistemas de representação escrita aplica-se também às disciplinas em que se reparte o conhecimento gramatical: os sons e o seu sistema de representação gráfica (Ortoépia e Ortografia) são pois prévios à gramática, imprescindíveis para proceder num terceiro momento à «analyse miuda do discurso», e para reduzir este a «certas classes, as partes elementares da oração», tomando em conta para isto não apenas os critérios semânticos habituais mas também outros formais e funcionais. Temos assim a disciplina chamada naquela altura *Etimologia* ou *Analogia*, a corresponder com a actual morfologia.

¹⁰ A citação de Cícero com que se abre o livro (“Usum loquendi populo concessi, scientiam mihi reservavi”, Cic. Orat. 48) é bem eloquente da importância e autoridade que se arroga o gramático.

¹¹ “Toda a grammatica é um systema methodico de regras, que resultam das observações feitas sobre os usos e factos das linguas” (Barbosa 1866: XI).

Como culminação do estudo gramatical vêm em quarto lugar as regras gerais e as possibilidades combinatórias que o uso faz das partes da oração «para exprimir todas as operações do entendimento», a sintaxe. Os primeiros autores que segundo Soares Barbosa trataram este tema na cultura clássica são Plato e Ennio e segundo ele sem se ocuparem propriamente da Gramática, mas da especulação sobre a gênese natural ou arbitrária do signo linguístico¹².

As pretensões lógicas desta gramática visíveis no título, na estrutura e em numerosas referências através das suas páginas tem a ver com a influência cultural e gramatical francesa, que desde a *Grammaire Générale et raisonnée* de Port-Royal acentuou ainda mais o pendor racionalista da gramática ocidental e que se deixou sentir nas tradições gramaticais do resto do continente, influxo que não se viu roto até à aparição revolucionária da corrente histórico-comparatista a meados do século XIX. Na tradição gramatical portuguesa Soares Barbosa fora já precedido neste logicismo por Manuel Coelho de Sousa, autor de um *Exame da Syntaxe e Reflexões sobre as suas Regras* (1729), por Verney no seu compêndio *Grammatica Latina Tratada por Hum Methodo Novo Claro e Fácil* e pela gramática filosófica de Melo e Bacelar (1783). Contudo, foi Soares Barbosa quem demonstrou na prática com a sua notável gramática as potencialidades – e também os limites – deste tipo de obras quando aplicados à realidade linguística particular do português.

Ensinar a língua e as suas leis era como dissemos equivalente a ensinar a pensar, porque linguagem e pensamento estavam em íntima ligação:

sendo a grammatica de qualquer lingua a primeira theoria que principia a desenvolver o embrião das idéas confusas da idade pueril; e dependendo da exactidão de seus princípios o bom progresso nos mais estudos, ella deve ser uma verdadeira logica, que ensinando-se a fallar, ensine ao mesmo tempo a discorrer. Que por isso a grammatica foi sempre reputada como uma parte da lógica, pela íntima conexão que as operações do nosso espirito tem com os signaes que as exprimem. E esta é a razão porque os antigos philosophos e os stoicos principalmente se faziam cargo d'ella nos seus tratados de philosophia, como Protagoras, Platão, Aristoteles, Theodectes, Diogenes, Chrysippo, Palemom e outros, sobre os quaes se pôde ver Laercio nas suas vidas, e Quintiliano Inst. Orat. I, 6 (Barbosa 1866: XII).

A gramática, guiada pela lógica, está em estreita ligação com as ciências, porque constitui como elas um instrumento analítico, e do mesmo modo que a linguagem formalizada por excelência, a matemática, tem o seu particular sistema de notação,

¹² Em Rousseau no seu ensaio sobre a origem das línguas encontram-se também estas quatro fases evolutivas (Rousseau 1980: 39-46). No artigo “Grammaire” da Enciclopédia aparece de modo mais resumido, falando-se de escrita simbólica substituída pela escrita ortográfica (Auroux 1973: 85-86).

a linguagem possui também o seu sistema, o da escrita alfabética, culminação do processo evolutivo antes explicado:

As linguas são uns methodos analyticos que Deus deu ao homem para desenvolver suas faculdades. Ellas dão o primeiro exemplo das regras da analyse, da combinação e do methodo, que as sciencias as mais exactas seguem nas suas operações [...]. Os progressos que com os algarismos fez a sciencia dos numeros dão a conhecer assás a importância também da escriptura alphabetica para os mais conhecimentos (Barbosa (1866: X).

Este logicismo era, porém, anterior, porque já estava presente em toda a tradição gramatical ocidental. É um facto bem estabelecido que as categorias gramaticais criadas pelos gregos e depois transmitidas a Roma derivaram-se das categorias habilitadas pelo pensamento filosófico para a análise da realidade e que noções como substância, accidente, accção, passividade, actividade, temporalidade..., foram transferidas da filosofia para a gramática, de modo que toda a tradição medieval, que continuara difundindo as gramáticas latinas através dos modelos de Donato e Prisciano ou a mesma tradição renascentista trabalha a partir destas noções e das classes de palavras que essas noções determinaram historicamente. Também não podemos esquecer o vínculo directíssimo estabelecido entre a *Grammaire* de Port-Royal e a famosa *Lógica* (1662) da mesma escola. Isto não significa, porém, que existisse uma completa unanimidade à hora de determinar quais as classes de palavras presentes em cada idioma, por não se utilizar apenas o critério semântico –embora tivesse este um papel primordial– mas também outros de carácter formal ou mesmo sintáctico ou distribucional e porque as particularidades linguísticas não eram sempre facilmente reduzíveis a um modelo gramatical único e universal.

Esta tendência lógica vê-se na gramática de Soares Barbosa em diferentes aspectos da descrição gramatical:

- a) No seu propósito de enunciar princípios e regras, procurando que estas fossem em número reduzido que a modo de leis se applicassem de maneira universal.
- b) No seu propósito de explicar as excepções às regras anteriores como falsos casos de excepção, como irregularidades aparentes. Aqui desempenhava um papel importante o procedimento da *elipse*, autêntico mecanismo que permitia interpretar *ad hoc* as irregularidades morfossintácticas como anomalias aparentes, sublinhando, pois, o carácter geral das leis da gramática. Este é um dos aspectos, por outra parte, que liga esta tradição gramatical logicista a modelos modernos como a da gramática generativo-transformacional, corrente que reivindicou os contributos cartesianos de Port-Royal num famoso livro de Noam Chomsky (1969).

5. Gramática universal vs. gramáticas particulares; gramática prática vs. gramática filosófica

A pluralidade das línguas do mundo constituía um facto importante que devia ser explicado por aqueles que afirmavam a unidade linguística da humanidade. A diversidade linguística seria para eles «aparente», porque às divergências na superfície corresponderia uma igualdade de fundo, por existirem uma série de princípios universais comuns: umas categorias lógico-semânticas «profundas» de carácter universal dirigiriam todas as línguas, a subjazer às particularidades e acidentes de superfície. São teses que, por outra parte, no século XX tomou para si e desenvolveu a semântica generativa¹³.

A parte lógica e racional da linguagem dependia das operações intelectuais, do pensamento. Se este era comum à humanidade, podia-se deduzir tanto a igualdade de fundo das línguas como a existência de uns princípios gramaticais de carácter geral:

as palavras [são] sinais artificiais das idéas e suas relações, e como taes sujeitos ás leis psychológicas que nossa alma segue no exercicio das suas operações e formação de seus pensamentos. As quaes leis sendo as mesmas em todos os homens de qualquer nação que sejam ou fossem, devem necessariamente comunicar ás línguas, pelas quaes se desenvolvem e exprimem estas operações, os mesmos princípios e regras gerais que as dirigem (Barbosa 1866: XI).

Quando as observações linguísticas vão referidas a um idioma concreto, a gramática será particular. Se as observações abrangem os usos e factos de todos ou a maior parte dos idiomas conhecidos, a sua gramática será geral. Se tomamos na sua literalidade esta afirmação de Soares Barbosa, esta gramática geral não seria, pois, apenas o resultado de um *a priori* filosófico, mas produto da observação de factos, neste caso da análise de diversas línguas, muitas delas com sensíveis diferenças estruturais. O problema será como estabelecer essa identidade de fundo das línguas a partir de uma variação formal tão alta.

Há uma outra divisão entre gramática prática e filosófica que responde às diferenças de público e de método didáctico: a gramática será «prática» ou «rudimentaria» quando dirigida ao conhecimento pragmático da língua, ao domínio das

¹³ Estas teses da identidade de fundo das línguas acabaram conduzindo a projectos de elaboração de uma língua universal, na esteira das tentativas de Raimundo Lúlio ou Leibniz. Em Portugal, foi José Maria Dantas, matemático, quem em 1800 publicou um projecto neste sentido na sua *Pasigraphia, composta e dedicada ao Serenissimo Senhor Infante D. Pedro Carlos* (veja-se Calafate 2001, III: 243). Sobre o projecto de língua universal ideado pelo padre Sarmiento, veja-se Lázaro Carreter 1985: 131-141.

regras básicas da sua pronúncia, escrita ou regras gramaticais, utilizando a memorização. Quando o objectivo seja conhecer as leis psicológicas e «as primeiras causas ou razões dos procedimentos uniformes que todas as línguas seguem na análise e enunciação do pensamento» (Barbosa 1866: XII) mover-nos-emos no terreno da gramática filosófica e científica. A gramática prática ou rudimentar, porém, deve fundar-se nos princípios da gramática filosófica, e não seria mais do que a sua aplicação aos níveis mais elementares do ensino.

O propósito veicular que tinham as gramáticas vernáculas, consistente em facilitar um posterior aprendizado da língua latina, ainda está presente¹⁴, e para justificar isto pode apoiar-se o nosso autor na autoridade de gramáticos anteriores como João de Barros.

6. Gramática contrastiva e tipologia linguística

As gramáticas racionalistas e filosóficas caracterizam-se pela sua perspectiva sincrónica. Todas elas são prévias à grande revolução que tanto na conceição das línguas e da ciência da linguagem como do ponto de vista metodológico vai supor a aparição da gramática histórica e comparativa do século XIX. É evidente, porém, que os gramáticos do XVIII conheciam com bastante exactidão a variação linguística e mesmo que não se podiam desinteressar dela, porque, defendendo a identidade essencial de todas as línguas, tinham que reconhecer e explicar o fenómeno das variação linguística. Por outro lado, sabemos que é nesta época quando surgem os primeiros ensaios comparatistas (Rousseau, Sarmiento...) e os primeiros projectos de elaboração de uma língua universal de carácter artificial (Lázaro Carreter 1985: 131-141).

A identidade essencial das línguas defendida por estes gramáticos não é um facto substancial, mas de tipo formal. Não importa, pois, nem a substância nem o processo de mudança que a conduz de uma língua à outra, não interessa pois a evolução diacrónica (por exemplo, do latim às diversas formas de romance), processo nunca explicado ou apresentado como consequência de um uso cuja mecânica e regras não se explicitam. O importante é a identidade de fundo entre os dois pontos, entre duas sincronias, a de partida e a de chegada –latim e português, por exemplo– em tanto que maneiras diferentes de ordenar e exprimir a mesma realidade conceptual da proposição.

Por isso, e embora na gramática de Soares Barbosa encontremos exemplos de comparatismo, nunca é com a função que este terá na linguística histórico-comparatista; os contrastes da estrutura linguística servem a Soares Barbosa para estabelecer

¹⁴ “Aquelles que aspiram a estudos maiores, e para entrarem n’elles tem de aprender as línguas sabias, levam uma grande vantagem com aprender primeiro a grammatica de sua lingua” (Barbosa 1866: XIV).

autênticas classes tipológicas de línguas, como se evidencia por exemplo no final do capítulo V da sua gramática, dedicado à preposição: depois de reduzir os advérbios a formas abreviadas da determinação com preposição, o nosso gramático classifica as línguas segundo a maneira em que se marca essa determinação: línguas a utilizarem um sistema preposicional (entre as que situa o português, espanhol, francês, mas também línguas semíticas como o hebreu); línguas a utilizarem um sistema de afixos ou posições chamados «casos» (o latim, mas também línguas como o vasco¹⁵ e as línguas indígenas do Peru); e uma terceira categoria constituída por aquelas que utilizam a combinação de preposições e caso (grego e latim, alemão e arménio). Para explicar a existência deste terceiro grupo a combinar os dois procedimentos anteriores recorre ao reduzido número de casos presentes nestas línguas, apenas seis, facto que obriga a que seja a combinação do sistema casual com o preposicional o que permita a expressão de muitas relações.

Esta classificação das línguas, seguindo critérios formais e funcionais, parece-nos notável, porque demonstra não a ignorância das observações e os factos que acabariam fundado o método comparatista, mas um interesse precisamente no contrário: no que de permanente e imutável tinha o fenómeno linguístico. Os gramáticos de tendência filosófica atendiam ao que de comum tinham as línguas, o seu olhar estava posto na gramática universal, e era a partir desta que se estabeleciam e se entendiam as gramáticas particulares e os seus procedimentos. Tratava-se, pois, de um modelo de convergência: era por isso que Soares Barbosa podia dar as equivalências funcionais de cada caso latino em português, acudindo a critérios como o uso de artigo, a posição posterior ao verbo, o uso de uma ou outra preposição etc. (Barbosa 1866: 230). A gramática histórica, diferentemente, privilegia o específico, e só num processo de explicação genética consegue chegar a uma primitiva unidade, que, por outra parte, não se pretende universal, mas parcial: línguas indogermânicas de um lado, semíticas do outro, etc. Trata-se, pois, de um modelo baseado na divergência, no diferencial.

Significativamente, esta corrente historico-comparatista abandonará, por inúteis e erróneas, as especulações sobre as origens da linguagem, tão caras aos defensores de uma gramática universal.

Por outro lado, as gramáticas filosóficas embora sejam sincrónicas no seu espírito não apresentam uns limites cronológicos definidos a determinarem as fronteiras temporais em que é aplicável a gramática: se em Soares Barbosa as referências à

¹⁵ Soares Barbosa cita a obra de Larramendi que devia conhecer bem: “O padre de Larramendi, jesuita que em 1729 deu á luz uma grammatica vasconga, escripta em hespanhol, debaixo do pomposo titulo de *El imposible vencido, ou Arte de la Lengua bascongada*, impresa em Salamanca, no cap. IX da II parte, reconhece que estas terminações, a que elle chama posições, semelhantes aos affixos hebraicos, equivalem ás preposições” (Barbosa 1866: 229).

língua trovadoresca marcam o que de diferencial e arcaico têm esta relativamente ao português do XVIII, não estão claras as fronteiras com outras épocas como demonstram, por exemplo, as citações de João de Barros ou Camões, ou as alusões a formas linguísticas reconhecidas como arcaicas pelo nosso gramático.

Também o carácter sincrónico da gramática não se contradiz com referências ocasionais à mudança linguística que vemos, por exemplo, quando Soares Barbosa fala do uso do participio perfeito activo, em concordância na língua antiga, para constatar que já na época clássica da língua portuguesa se produziam os dois usos (Barbosa 1866: 193). Outro exemplo de referência diacrónica temo-lo quando explica o gramático como na «linguagem antiga» (a etapa galego-portuguesa), a desinência de segunda pessoa de plural era /-des/ e não /-is/, forma a primeira de que ficariam rastros em formas verbais como o infinitivo pessoal e futuro imperfeito de conjuntivo (*amardes, fordes*). Não há no exemplo, porém, historicismo, mas a simples constatação de duas formas diferentes, correspondentes a dois estados da língua, sem qualquer interesse em fixar os processos de evolução fónica que conduziram da primeira forma para a segunda. Por outro lado, a facto de fazer isto para «facilitar a intelligencia dos manuscriptos mais antigos» (Barbosa 1866: 168) prova que as gramáticas, à parte do seu valor descritivo, pretendiam também ser instrumento auxiliar para aqueles que deviam manejar documentação antiga.

7. Influências

A primeira e mais visível das influências é a dos gramáticos franceses desde Port-Royal. Condillac e o seu sensualismo, influídos pelo empirismo de Locke, está muito presente¹⁶. O ser humano comunica-se com a realidade pelos sentidos e para Condillac as línguas foram construídas a partir de uma linguagem natural de gestos, que acabou por perder o seu carácter natural, e que através de um processo de estilização e abstração crescente deu lugar à linguagem abstracta. Também o trabalho de teorização e sistematização linguística seguido por Du Marsais¹⁷ e Beauzée, principais autores dos artigos linguísticos da *Enciclopédia* francesa (1751-1780) não deixou de ter a sua influência em Portugal, nomeadamente no Dicionário publicado pela Academia em 1793 (Hassler 2001: 119).

Soares Barbosa devia ser um bom conhecedor dos modelos gramaticais aplicados a diferentes línguas europeas e ele próprio confessa-se devedor das gramáticas que Arnaud, Wallis e Starris e Lancelot elaboraram para o francês, inglês e espanhol e italiano (Barbosa 1866:XIII). Está também logicamente presente a tradição grama-

¹⁶ Parte da obra de Condillac fora traduzida para português. Veja-se Hassler 2001: 107.

¹⁷ Sobre a contribuição de Du Marsais à *Enciclopédia* veja-se Du Marsais 1987.

tical hispânica: João de Barros, Contador de Argote, Amaro de Roboredo, Bento Pereira, Pedro José da Fonseca, ou a gramática da Real Academia Espanhola de 1771, embora os critérios do nosso autor sejam bastante diferentes dos utilizados pelos anteriores gramáticos portugueses.

A tradição gramatical portuguesa anterior e contemporânea é avaliada por Soares Barbosa a partir dos seus postulados logicistas e dos seus critérios sobre uma gramática pedagogicamente aceitável. Os seus juízos não são, porém, especialmente abundantes, citando ocasionalmente os *Rudimentos*, a obra de Amaro de Roboredo, Bento Pereira, João de Barros, Contador de Argote ou Reis Lobato. Esse escasso interesse em basear a sua doutrina gramatical na tradição portuguesa responde, na nossa opinião, não à ignorância dessa tradição, mas à sua convicção de ser a sua gramática, à qual aplica o qualificativo de «filosófica», a única com pretensões rupturistas, a que no campo da língua portuguesa se ajusta fielmente aos princípios de reforma propostos.

O outro motivo de distância com as gramáticas portuguesas anteriores ou coetâneas radica nos seus defeitos: censura que o sistema analógico usado nelas consista basicamente numa cópia servil da gramática latina, e que deixem assim inexplicadas «muitas observações necessárias sobre o genio particular e caracter da Língua Portuguesa» (Barbosa 1866: XIII) . Fala elogiosamente, porém, de Pedro José da Fonseca, autor dos *Rudimentos da Grammatica Portuguesa*, obra impressa em Lisboa em 1799, aplaudido por utilizar como modelo a *Gramática da Língua Castelhana* composta pela Real Academia Espanhola, «a qual entre as das línguas vulgares tem merecido um distinto louvor» (Barbosa 1866: XIV). Mas mesmo esta obra elogiada não está livre de defeitos: o primeiro é uma demasia de exemplos e regras, pois Soares Barbosa é partidário de abreviar e reduzir a gramática a poucos princípios de carácter geral. O segundo é que deixasse fora da Gramática ortoépia e ortografia, «partes essenciaes e importantes a qualquer grammatica vulgar».

Soares Barbosa orgulha-se também da importância e precocidade da tradição gramatical portuguesa do XVI, como podemos ver nas suas referências à gramática de Fernão de Oliveira:

Portugal conheceu grammaticas portuguezas ainda antes que outras nações civilisadas tivessem uma na sua lingua. Quando Ramos em 1572 publicou a primeira grammatica da lingua franceza, já Portugal tinha a de Fernão d’Oliveira dada á luz em 1536, e a de João de Barros em 1539 (Barbosa 1866: XIII).

Parece evidente que esta apologia da gramática do XVI tem um carácter compensatório, que se trata em realidade de atenuar que essa precocidade gramatical por-

tuguesa veio seguida de um período de decadência e estagnação. As glórias do passado contrabalançam assim a decadência do presente. A sua gramática pretenderá por isso modernizar a tradição gramatical portuguesa e levá-la ao nível de outras tradições gramaticais, nomeadamente a francesa.

A gramática espanhola aparece também citada em diferentes ocasiões (Gramática da Real Academia), o qual demonstra que o nosso autor manejava abundantes fontes que lhe serviam como modelos a que aproximar ou de que afastar a sua. Na *Introdução*, Soares Barbosa elogia também entre as figuras do passado a do gramático Sánchez (Sánchez de las Brozas) nomeadamente pelo carácter anticonvencional, rupturista e logicista das teses do gramático salmanticense¹⁸.

Também é possível que, naqueles casos em que Soares Barbosa fez referência ao vasconço para comparar as diferenças existentes entre o uso dum sistema preposicional e de outro consistente num sistema sufixal, tivesse como fonte a obra de Larramendi, amplamente difundida na altura¹⁹.

8. A gramática: caracterização e partes

Como já vimos são quatro: duas a ocuparem-se do aspecto «mecânico» da língua: a Ortografia e a Ortoépia. As outras duas a tratarem da parte «lógica» das língua: a Etimologia (isto é, a Morfologia) e a Sintaxe.

a) Ortoépia

Compreende duas partes:

- O conhecimento dos sons fundamentais da língua. Seria estritamente este o campo de estudo da ortoépia.
- «As modificações musicas de que os mesmos são susceptíveis» (canto, melodia, acentos, compasso e ritmo, quantidade). Esta segunda parte recebe o nome de Prosódia. Segundo Soares Barbosa, foi esta posterior à ortoépia, pois surgiu a sua necessidade quando os estrangeiros começaram a estudar outras línguas diferentes da sua. Os acentos da língua grega, posteriores às letras do seu alfabeto, assim como o facto de os gregos começarem a estudar a prosódia em época bastante tardia (nos tempos de Cícero) seria para ele prova disto. Eis a causa de escolher este termo de *Ortoépia* para designar conjuntamente a Ortoépia em sentido estrito e a Prosódia.

¹⁸ “Mas felizmente aconteceu em nossos tempos, que Sanches principiasse entre os hespanhoes a sacudir o jugo da auctoridade e preocupação n’estas materias, e introduzindo na grammatica latina as luzes da philosophia, descobrisse as verdadeiras causas d’esta lingua” (Barbosa 1866: XIII).

¹⁹ Sobre a figura de Larramendi e a filologia hispânica do XVIII veja-se Lázaro Carreter 1985: 112-113.

b) Ortografia

Incluem-se nela os signos gráficos que representam os sons da língua e um estudo dos diferentes sistemas ortográficos existentes. Vista a inexistência de uma ortografia regular e aceite por todos, diferencia entre as ortografias «etimológica», «usual» e «de pronúnciação»; esta última oferece-se como uma revolucionária proposta de reforma ortográfica coincidente em muitos pontos com a de Verney. A «usual» não é mais do que a ortografia consagrada pelo uso na altura, uma ortografia etimológica aplicada de modo incoerente e por isso particularmente censurável:

Entre estas duas orthographias caminha a usual. Assim chamada, porque não tem outra auctoridade senão a do uso presente e dominante, já para seguir as etymologias e introduzir arbitrariamente escripturas mui alheias da pronúnciação presente; já para não fazer caso da derivação mesma, e incoerente em seus procedimentos escrever, p. ex.: *he*, *huma*, com *h*, que não ha na origem latina, e *filosofia* e *fysica* com *f* e *z* que não ha nas palavras gregas (Barbosa 1866: 39)

Vai a seguir um apartado destinado ao estudo dos signos de pontuação. Se a ortografia aparece ligada à representação da palavra individual, a pontuação articula a união das palavras no discurso. O seu sistema de pontuação está muito bem articulado, é muito mais rígido do que o proposto por Verney e responde à divisão progressiva do discurso em partes menores:

As primeiras subdivisões parciaes das orações, ou juizos que fazem parte de outros, são marcadas pelas *vírgulas*, que é a pontuação mais fraca e inferior.

A segunda divisão do periodo em membros ou proposições totaes, que contém em si as primeiras subdivisões parciaes, é marcada com *ponto e vírgula*.

E a divisão mestra, ou principal das duas partes de qualquer periodo, antecedente e consequente, que comprehendem em si todas as outras subdivisões subalternas, é marcada com *dois pontos*.(Barbosa 1866: 61)

c) Etimologia

É a primeira das duas partes lógicas ou racionais da linguagem. As classes de palavras são concebidas a partir de critérios primeiramente semânticos e secundariamente formais e funcionais:

Esta analyse do discurso dependia de muitas observações particulares e de muitas combinações para d'ellas se formarem noções geraes, que reduzissem a certas classes as partes elementares da oração segundo as suas significações e analogias, e bem assim a regras geraes as varias combinações que o uso fazia das mesmas (Barbosa 1866: IX).

A analogia e unidade formal, a pesar da diversidade e das irregularidades aparentes, são o objecto da etimologia (mas não apenas dela, porque esta insistência em sublinhar as regularidades do funcionamento lingüístico aparece também na Sintaxe). A função da gramática estaria em «descobrir todas as analogias de uma língua particular, e reduzir a ellas muitas anomalias que os ignorantes contam por taes, não o sendo realmente» (Barbosa 1866: XII).

Embora o estudo morfológico preceda no texto gramatical o estudo sintáctico, e podendo efectivamente estabelecer classes de palavras segundo critérios semânticos e morfológicos, não está absolutamente clara esta anterioridade do ponto de vista lógico, por falar o próprio Soares Barbosa das classes de palavras como “partes da oração”: isto é, antes de definir as partes da oração realmente existentes estabelece o gramático um marco superior, o princípio da inclusão dessas classes de palavras numa unidade maior chamada proposição e em cujo interior identificar-se-iam essas unidades morfológicas.

Soares Barbosa, porém, procede na sua exposição de menor a maior e afirma portanto a precedência da morfologia:

os homens se occupariam em considerar primeiro as palavras que são signaes assim das idéas que fazem o objecto dos nossos pensamentos como das relações que as mesmas podem ter consigo e com outras, do que em considerar estas mesmas palavras combinadas e coordenadas entre si em ordem a exprimirem o pensamento (Barbosa (1866: X).

As classes de palavras podem exprimir ideias (substantivos, adjectivos...) ou relações (verbos, preposições, conjunções...). Como todas as gramáticas filosóficas a de Soares Barbosa divide os verbos em duas grandes classes, a do «verbo substantivo» (formada apenas por um verbo, *ser*) e a dos «verbos adjectivos» (todos os outros, sempre analisáveis como união de uma forma do verbo substantivo e um adjectivo: *amar = ser amante*). São notáveis e de muito interesse os seus intentos de ordenação das formas adjectivas (capítulo III) e das preposições (capítulo V).

d) Sintaxe

As palavras unidas no discurso são o objecto da Sintaxe. A sintaxe de Soares Barbosa e da época, diferentemente à nossa, não se ocupava das funções oracionais, nem considerava a existência de unidades intermédias entre o nível da palavra e o da oração, mas tratava da relação de *concordância*, existente entre algumas classes de palavras quando agrupadas em unidades maiores, como a do substantivo com o adjectivo ou a do sujeito com a forma verbal, ou da *regência* (formas adjectivas ou verbais complementadas com a ajuda de uma determina-

da preposição). Estas noções de concordância e regência eram herança directa da tradição latina. A noção de função sintáctica e de complemento, próprias da gramática moderna, estavam já presentes, como revela o facto de falar de verbos transitivos ou intransitivos ou de complemento terminativo, por exemplo. A ordenação dos elementos na oração também é tratada de modo minucioso como se vê nos capítulos destinados à «construção direita» e à «construção invertida», centrada esta última na análise da anástrofe e do hipérbato.

9. Ideal linguístico vs. usos estabelecidos

Algumas questões da gramática de Soares Barbosa põem à luz a complexidade resultante do confronto entre os modelos teóricos que o gramático como possuidor do saber linguístico propõe e a realidade teimosa de uma prática que não se ajusta a eles e que por vezes é manifestamente contrária. É aqui, na distância entre uma teoria linguística que se pretende prescritiva e as resistências a esses intentos de modelação por parte de instituições e práticas tradicionais em terrenos como as convenções ortográficas ou os modelos sintácticos da linguagem literária que podemos observar os limites reais da autoridade do gramático sobre a linguagem.

A questão ortográfica é a primeira delas. No capítulo destinado à ortografia, Soares Barbosa elabora uma proposta radical de simplificação ortográfica, que encontra uns claros antecedentes na primeira carta de Verney no *Verdadeiro método de estudar*. As regras dessa nova «ortografia de pronúnciação» reduzem-se a uma muito simples:

Qualquer palavra que se queira escrever, pronuncie-se primeiro bem, e distinguidos todos os sons de que é composta, estes se escrevam pela mesma ordem, com os caracteres que lhes competem nos abecedarios completos e exactos /.../ e a palavra assim escrita ficará sem erro de ortografia (Barbosa 1866: 52).

A reforma supunha eliminar <h> salvo nos dígrafos <ch, lh, nh>, simplificar a representação das sibilantes (<s> para representar em todos os casos /s/, <z> para representar o fonema /z/, <j> para representar em todos os casos a palatal fricativa sonora, e a dupla opção de <ch> e <x> para a fricativa palatal surda), eliminar as geminadas e as combinações consonânticas dos grupos cultos, etc.

O programa ortográfico de Soares Barbosa, radical nessa altura, ficou no plano da pura teoria desde o momento que o próprio autor apenas utilizou essa nova ortografia chamada «de pronúnciação» nas seis páginas do capítulo em que foi exposta (Barbosa 1866: 52-57). Uma ortografia pois que não teve qualquer consequência prática na sua época e na própria obra de Soares Barbosa. Em qualquer caso, na

história da ortografia portuguesa, o contributo teórico tanto de Soares Barbosa como de Verney - com sugestões de reforma muito similares e que talvez inspiraram o primeiro - é inegável e sinala a direção de futuros planos de reforma no XIX (Kemmler 2001: 253-266).

É na sintaxe que vemos novamente este contraste entre teoria e prática. A gramática de Soares Barbosa procura ir como outras muitas acompanhada de citações tiradas daqueles autores que exemplificam e canonizam o usos correctos da língua portuguesa, mais paradoxalmente os fragmentos de prosa ou poesia que fornece o autor são em muitos casos censurados pelo próprio gramático: por terem uma construção sintáctica pouco clara ou por uma disposição dos elementos oracionais marcada pelo hipérbato, a impedir este o reconhecimento do sentido do texto .

O princípio teórico da claridade e da simplicidade e da construção guiada pelo critério da univocidade do significado proposto pelo gramático entrava em confronto, pois, com uma tradição literária portuguesa derivada de práticas e tradições retóricas muito diferentes, manifestas numa sintaxe barroca e deliberadamente complexa e no gosto pela ambiguidade e os jogos de palavras. É por isso que muitos autores portugueses aportados como exemplos viram-se em autênticos contra-exemplos, modelos em negativo objecto da repreensão do gramático. O rígido logicismo aplicado por Soares Barbosa nos seus modelos sintácticos não deixava perceber que além da lógica outros motivos de carácter psicológico e estético actuavam também nas escolhas sintácticas. Desde o seu saber gramatical racionalista e de influência francesa, a avaliação dos usos sintácticos tirados da literatura portuguesa clássica e contemporânea não podia deixar de ser negativa.

Referências bibliográficas

- Assunção, C. da Costa (1997): *Para uma gramaticologia portuguesa. Dos primórdios do gramaticalismo em Portugal a Reis Lobato* (Vila Real: Universidade de Trás-Os-Montes e Alto Douro).
- Assunção, C. da Costa (2000): *A Arte da Grammatica da Lingua Portuguesa de António José dos Reis Lobato*. Estudo, edição crítica, manuscritos e textos subsidiários (Lisboa: UTAD –FCT).
- Auroux, S.(1973): *L'Encyclopédie. "Grammaire" et "Langue" au XVIIIe. siècle* (France: Mame).
- Bacelar, B. de Lima e Melo (1996) [1783]: *Gramática Filosofica da Língua Portuguesa por Bernardo de Lima e Melo Bacelar. Reprodução facsimilada da edição de 1783 com introdução e notas pelo académico correspondente Amadeu Torres* (Lisboa: Academia das Ciências).

- Barbosa, J. Soares (1866) [1822]: *Grammatica Philosophica da Lingua Portuguesa ou Principios da Grammatica Geral applicados á nossa linguagem* (Lisboa: Typographia da Academia Real das Sciencias).
- Calafate, P. (dir.) (2001): *História do pensamento filosófico português*. III. *As Luzes* (Lisboa: Caminho).
- Casteleiro, J. Malaca (1980/81): “Jerónimo Soares Barbosa: um gramático racionalista do século XVIII”, *Boletim de Filologia*, XXVI: 101-110.
- Chomsky, N. (1967): *Linguística cartesiana* (Madrid: Gredos).
- Du Marsais, C. Chesneau (1987): *Les véritables principes de la grammaire et autres textes. 1729-1756* (Paris: Librairie Arthème Fayard).
- Hassler, G. (2001): “O problema da significação das palavras e as teorias linguísticas dos séculos XVII e XVIII”, em Thielemann, W. (ed.): *Século das Luzes- Século de Pombal*: 99-127 (Frankfurt am Main: TFM).
- Kemmler, R. (2001): “Para uma história da ortografia portuguesa: o texto metaortográfico e a sua periodização do século XVI ate à reforma ortográfica de 1911”, *Lusorama*, 47-48: 128-319.
- Lázaro Carreter, F. (1985) [1949]: *Las ideas lingüísticas en España durante el siglo XVIII* (Barcelona: Crítica)
- Mariño Paz, R. (1998): *Historia da lingua galega* (Santiago de Compostela: Sotelo Blanco).
- Monteagudo, H. (1999): *Historia social da lingua galega* (Vigo: Galaxia).
- Rousseau, J. J. (1980) [1761]: *Ensayo sobre el origen de las lenguas* (Madrid: Akal).
- Verney, L. A. (1949) [1746]: *Verdadeiro método de estudar* (Lisboa: Sá da Costa).

‘Línguas em contacto’ na Galiza Baixo-medieval¹

Júlio Diéguez González
Universidade de Santiago de Compostela

Resumo:

O principal elemento legitimador dos estados *nacionais* é actualmente a língua, como foi a religião outrora. A língua tem na nossa época o mesmo papel que a religião na sociedade medieval: é um elemento basilar na legitimação e no funcionamento da instituição estatal. Há uma diferença radical num ponto concreto: na Idade Média não parece ter havido rejeição à mistura de línguas, o que é uma característica constante nas modernas sociedades industriais. Neste artigo estudam-se alguns factores que terão dificultado a substituição do galego-português pelo castelhano na elaboração dos documentos jurídicos na Galiza dos finais da Idade Média: (i) o peso de uma tradição de vários séculos, e (ii), ser o galego-português a língua da oralidade e a mais inteligível para as partes intervenientes nos actos jurídicos. Desde o século XIII Castela representa o centro do poder, e há um relacionamento intenso da nobreza galega com a castelhana. Nos finais do século XV muitos nobres são obrigados a residir na Corte. Nessa época o castelhano já era percebido como a língua da autoridade.

Palabras chave:

Contacto de línguas. Conflito linguístico. Galego-português medieval.

Abstract:

Just like religion before it, language in our day has become the principal legitimising element for national states. Today, language plays the same role as religion did in medieval society: it constitutes a basic element in the legitimisation and functioning of the state as an institution. Yet there is a radical difference in one aspect: during the Middle Ages there does not appear to have been a sense of rejection of the mixing of languages, which is, however, a constant feature of modern industrialised societies. This article looks at a number of factors which have led to difficulties in replacing Galician-Portuguese for Castilian Spanish in legal documents in the late Middle Ages: (i) the weight of a tradition dating back several centuries and (ii) the fact that Galician-Portuguese was used for oral communication and was therefore the most intelligible language for all those parties intervening in legal proce-

¹ Este estudo insere-se no projecto de investigação “Prosa historiográfica e documental da Galiza medieval: séculos XIV-XV” subsidiado pela Dirección Xeral de Universidades e Investigación da Consellería de Educación e Ordenación Universitaria da Xunta de Galicia (PGIDT00PXI20404PR). Siglas: A. H. N.= Archivo Histórico Nacional (Madrid); A. H. P. O.= Arquivo Histórico Provincial de Ourense; A. C. O.= Arquivo da Catedral de Ourense; A. A.= Arquivo do Mosteiro de Benedictinas de Sam Paio de Anteaaltres (Santiago de Compostela); A.H.U.S.=Arquivo Histórico da Universidade de Santiago. A. F. P.=Arquivo do Mosteiro de Ferreira de Pantom.

dures. Since the 13th century many members of the aristocracy were forced to reside in Court. By this time Castilian Spanish was already held to be language of “authority”.

Key words:

Contact between languages. Linguistic conflict. Medieval Galician-Portuguese.

A língua é, nos nossos dias, o principal elemento legitimador dos estados *nacionais*, como foi a religião outrora. Do mesmo modo que o processo de *secularização* conduziu a religião para o âmbito da privacidade dos cidadãos, não sem antes ter deixado *marcada* fortemente a institucionalidade dos estados *ex-confessionais* em questões como o calendário, etc., cabe pensar que relativamente à língua pode no futuro dar-se uma evolução idêntica, e talvez possam surgir estados linguisticamente *não confessionais*, o que provavelmente já fará necessário um conceito e um termo diferente do actual: “estado”.

Os estados ocidentais actuais são neutrais em matéria religiosa em virtude do facto de fundamentar a sua legitimação em valores e conceitos *seculares*, distanciados das crenças das comunidades religiosas que coexistem nas sociedades correspondentes. Essa neutralidade teria sido impossível na época medieval, em que o poder dos monarcas, no sentir da população cristã ou muçulmana das monarquias peninsulares, por exemplo, derivava directamente da vontade de Deus ou de Alá, o que lhe outorgava a responsabilidade primária de proteger e preservar os interesses da comunidade religiosa que justificava a existência do reino. Do mesmo modo, é impossível imaginar na nossa época um teórico estado linguisticamente neutro em que várias comunidades linguísticas possam coexistir em termos de perfeita equidistância dos poderes institucionais sobre a base da consideração da língua como *prática estritamente privada* em que os poderes públicos não podem nem devem interferir. Essa impossibilidade deriva do facto de estarem as comunidades estatais vinculadas por textos jurídicos, imperativos para todos os cidadãos sujeitos ao cumprimento da lei, textos que por força têm de ser redigidos numa língua natural cuja fixação escrita é premissa indispensável para a existência de qualquer entidade estatal. Só poderíamos imaginar um estado linguisticamente não confessional se fosse possível redigir uma constituição e institucionalizar um ordenamento jurídico completo em linguagem matemática, por exemplo. Então sim, as línguas naturais poderiam ser secularizadas, reduzidas à mesma dimensão que têm nos nossos dias nas sociedades ocidentais as práticas religiosas².

² Não é que no topo do organograma social nas sociedades actuais esteja um texto (um pacto, de facto), e nas sociedades medievais um “ser supremo”; não, porque também nestas há um texto no topo da organização social, como é a Bíblia ou o Alcorão. A diferença é que na Idade Média trata-se de textos sacros, revelados, imutáveis e absolutos, enquanto na época actual são pactuados, susceptíveis, portanto, de modificação. Mas actualmente o exercício pleno dos direitos pelos cidadãos exige o conhecimento dos textos legislativos, e a participação social obriga-os a exprimirem-se na língua que os veicula, o que não ocorria de maneira nenhuma na época medieval.

Parece claro que a rejeição pela mistura de línguas é, essencialmente, uma consequência do papel que cabe actualmente a estas na organização, arquitectura e legitimação do poder estatal.

Falar em consciência de *identidade nacional* na época medieval seria totalmente anacrónico. Embora a ideia de *nação*, no sentido actual do termo, seja um conceito moderno, é todavia possível detectar na sociedade medieval os elementos indiciários da incipiente formação da consciência de identidade linguístico-cultural nas sociedades ibéricas. O início da escrita institucional³ em castelhano ou galego-português constituem mudanças qualitativas, pontos de inflexão que, evidentemente, não se podem considerar *consequência* ou *testemunho* do nascimento de uma consciência de identidade linguístico-cultural, mas como verdadeira *causa remota*. A progressiva institucionalização dos processos de interacção social criam modelos de comportamento que progressivamente vão sendo assumidos como *obrigatórios*, definitórios da *pertença* de cada sujeito a uma *colectividade*. No ponto de partida não parece que se possa considerar a existência de comunidades culturais baseadas na diversidade dos falares romances: não existe a comunidade galego-portuguesa face à castelhana ou leonesa; isto constata-se pelas notícias que temos dos comportamentos individuais no que diz respeito à *admissão* e à lealdade *comunitária*: a mudança de comunidade religiosa dos sujeitos dá lugar à mudança da denominação, é uma verdadeira *conversão* no que atinge à sua adscrição comunitária e ao código de lealdades. Essa admissão não existe quanto às colectividades linguístico-culturais: não há *naturalização* ou *nativização* castelhana ou galego-portuguesa dos sujeitos na sua mudança de partido político ou de residência. Nesse sentido a onomástica das pessoas fornece elementos indicativos da presença de uma consciência comunitária na mentalidade *popular* e na *auto-percepção* individual subjectiva. A Idade Média peninsular, como é notório, é caracterizada pela vigência de três grandes comunidades culturais correspondentes a três religiões, e o nome pessoal era indicador e ao mesmo tempo elemento de integração e socialização que vinculava (subjectiva e objectivamente) o sujeito a uma *religião-cultura*. As variantes linguísticas dos nomes pessoais, dos apelidos ou dos topónimos não terão sido percebidas como *opostas* no sentido *identitário grupal*, mas como perfeitamente equivalentes, relativas apenas aos costumes locais; a rivalidade entre os diferentes modos de falar, vestir ou mesmo dançar ou dar os pêsames ou os parabéns que podiam dar-se entre castelhanos e portugueses terão sido observados na mesma perspectiva que as diferenças locais entre diferentes terras de Castela, ou entre castelhanos e leoneses, ou entre castelhanos e andaluzes.

As guerras entre Castela e Portugal eram conflitos dinásticos, em boa medida equiparáveis aos que surgiam entre diferentes partidos nobiliárquicos. Não há uma

³ Referimo-nos sobretudo à elaboração dos textos legislativos e dos documentos jurídicos.

consciência comunitária directamente vinculada ao uso linguístico: a nobreza e os seus casamentos e alianças são prova da indiferença face à progressiva *normalização* linguístico-cultural que surgia da *padronização* e institucionalização progressiva das diferentes tradições culturais. A lírica cortesã do *Cancioneiro Geral* de Garcia de Resende, da segunda metade do século XV, e o facto de muitos portugueses terem trocado em castelhano, é indicador da nula capacidade coerciva da *consciência comunitária* no sentido de *fidelidade* ou *lealdade linguística*⁴. Cabe supor que a nobreza e as cortes dos monarcas possam ter sido de início resistentes à penetração deste tipo de *consciência identitária*, mas só até um dado momento: aquele em que se torna necessário obter o apoio dos populares nos conflitos territoriais que eclodiram e acabaram defrontando os diferentes reinos. Em Portugal, por exemplo, o apoio do povo à causa do Mestre de Avis pode interpretar-se, por uma parte, como resultado da *instrumentalização* por um grupo nobiliárquico-dinástico dessa *percepção* popular, mas ao mesmo tempo teve como consequência inevitável a *aceitação* institucional dessa perspectiva: é uma derrota da *consciência estamental* da nobreza face à *consciência de tradição cultural* dos sectores populares⁵. Nos dias de hoje a *lealdade linguística* é exigível de forma patente: um cidadão português não pode usar o castelhano na vida quotidiana residindo em Portugal, como é impossível para um espanhol usar regularmente o português em Castela por simpatia linguístico-cultural com o povo lusitano, ou que um autor literário possa praticar um bilinguismo como os do *Cancioneiro Geral* de Garcia de Resende.

O processo linear se deu em Castela e Portugal não se desenvolveu do mesmo modo na Galiza, onde se dá uma situação paradoxal: a institucionalização castelhana de uma tradição cultural galego-portuguesa. Esta interpretação justifica-se sobretudo pelo facto de que o início da escrita notarial em galego-português na Galiza, anterior à portuguesa, parece responder à prática de institucionalização do romance iniciada antes em Castela. Quando na Galiza surge a escrita em galego-português já há tempo que o território está inserido num espaço político em que convivem diferentes comunidades linguísticas. Posteriormente assenta-se uma comunidade de língua castelhana, numa posição de privilégio tanto em termos políticos como económicos. Esta força o aparecimento de uma situação de bilinguismo, em que só o castelhano ostenta uso institucional, impedindo a *institucio-*

4 Todavia existem testemunhos da mudança de *percepção* que se estava a produzir: são importantes a este respeito as informações das *Crónicas* de Fernão Lopes sobre as diferentes *tradições* que vigoravam em Castela e Portugal e como geravam e multiplicavam a percepção de *distância cultural* nos sectores populares.

5 Na medida em que os conflitos políticos forçam os monarcas a ganhar legitimidade no *sentir popular* são obrigados a reforçar os laços de *simpatia* cedendo no sentido da *institucionalização* cada vez maior da tradição cultural com que o povo se identifica. Surge assim um código robusto e rígido que obriga os sujeitos a adaptar a sua prática à *adscrição* comunitária linguístico-cultural.

nalização alternativa sobre a base da tradição cultural autóctone, que só tardiamente se produzirá⁶.

A imposição do nome deve ter sido e continua a ser o principal rito de *admissão* do sujeito para a comunidade social; por isso faz sentido considerar como comunidade *principal* na percepção colectiva da Idade Média a que vinha determinada pela tradição religiosa: porque a mudança de comunidade, a conversão, supunha uma mudança de nome, portanto uma nova *admissão* do sujeito no colectivo humano. A comunidade religiosa terá sido progressivamente substituída pela linguístico-cultural que, modernamente, se torna *nacional* em muitos âmbitos; não assim na Galiza, cuja situação actual é conhecida: à medida que a comunidade de prática linguística galego-portuguesa reforça e institucionaliza a sua *tradição linguístico-cultural* cresce a capacidade coerciva (a *pressão*) sobre os sujeitos de ambas as comunidades no que diz respeito à *lealdade linguística comunitária* nas práticas orais, e tem lugar um processo lento de *segregação* em que uma comunidade maioritariamente *bilingue* se cinde em duas cada vez mais *coesas* e mais tendencialmente monolíngues. Os nomes de pessoa são reveladores de vários aspectos dos processos ideológicos subjacentes: o nome pessoal marca actualmente cada vez mais a pertença à comunidade linguística da *língua autóctone*, embora não seja um elemento directamente gerador de lealdade linguística, mas muitas vezes *compensatório*. A família continua a ser um elemento basilar da *admissão linguística*; as práticas bilingues tradicionais (o castelhano como língua apenas institucional para um grande número de sujeitos) viram-se alteradas pela tendência para os reforços da lealdade linguística, sendo muitos os pais que *expulsam* os filhos da comunidade da *língua autóctone* para a castelhana, sobretudo para facilitar a integração e o sucesso no sistema educativo, fortemente ancorado na institucionalização castelhana.

Falar em castelhanização na Galiza relativamente à época medieval não deixa de ser um anacronismo porque, tanto quanto sabemos, não deve ter existido nessa altura uma consciência da individualidade dos diferentes romances como para que alguém pudesse pensar em termos de *língua castelhana* ou influência linguística de Castela do modo que hoje entendemos⁷; parece, todavia, que o processo deve

⁶ O galego-português da Galiza vive em simultâneo a sua eclosão o seu declínio: ao mesmo tempo que afirma a sua independência do latim já está iniciado o processo que o levará a perdê-la face ao castelhano. Como o movimento de reivindicação da língua própria que surge no século XIX nunca, até hoje, a reconheceu como *mista*, como nunca foi aceite a legitimidade do castelhano como elemento *genético* constitutivo —como se reconhece a do latim— surgiram duas maneiras contraditórias de encarar o problema: negando a presença do elemento castelhano na língua própria (nada que o falante diz espontaneamente é castelhano se não for sentido como tal, ou se intencionalmente não é; o castelhano da Galiza é um dialecto da língua própria) ou recusando-se a assumir como língua própria qualquer elemento da oralidade que esta tenha incorporado ou gerado a partir do castelhano, inclusive quando se afaste da norma ou do uso vigente nessa língua.

⁷ Estamos a referir-nos ao século XIII; para o século XV há evidência de ter existido um conceito perfeitamente claro do que era a “língua castelhana”.

ter sido percebido de modo muito mais difuso, sobretudo nos períodos mais antigos; há que ter presente que a homogeneidade dos falares romances era menor, o grau de elaboração das respectivas normas muito incipiente e relativamente às variedades dialectais pode não ter existido uma consciência de *status* inferior (sociológica e culturalmente) como em épocas mais recentes. Por isso, parece lógico pensar que, nos momentos em que a nobreza oriunda do âmbito leonês esteve na posse dos seus domínios territoriais galegos, pode ter sido um factor de castelhanização objectivo sem sê-lo em termos estritamente diatópico-linguísticos: faltando a consciência de uma identidade linguística *leonesa*, o elemento leonês deve ter passado a integrar-se, em termos de *consciência idiomática* dentro do que se teria considerado *falar castelhano*. Se isto ocorreu, como supomos ter ocorrido, a consequência inevitável é que a variabilidade interna do que era percebido como elemento *castelhano* se terá tornado muito maior⁸. Pode ser útil neste sentido a comparação com situações actuais: um linguísta será capaz, em abstracto, de discernir na *fala* castelhana de um nativo da Galiza o que é fonética galega, léxico castelhano, sintaxe *híbrida*, etc., mas essa análise não é operativa na mente do falante: *pensa* que fala castelhano, e é isso só que conta; na sua prática linguística está a ser um factor de castelhanização linguística duplamente eficaz: porque difunde o sistema linguístico castelhano obrigando outros utentes a adquirir competência passiva, e depois activa, e porque, ao acomodar o sistema castelhano a uma fonética diferente, está a alargar o seu campo de variabilidade e a facilitar a integração no mesmo de falantes com competência ainda menor. Se chegasse a perceber a distância entre a sua prática e *modelo* que crê estar a usar, modificará a prática no sentido do modelo, por isso o que verdadeiramente importa na escolha linguística é o que o sujeito *pensa* que fala, não a componente filológico-linguística da língua que usa⁹.

⁸ Desde o momento em que o utente que *pensa* que está a falar castelhano o faz sem renunciar a traços linguísticos *dialectais* e o contraste com a *língua da corte* não é suficiente como para que se sinta ridículo no seu modo de falar, o efeito da sua fidelidade linguística será, imediatamente, o alargamento do campo de variabilidade do que se considera *falar castelhano*.

⁹ Não significa, evidentemente, que o falante *aloglota*, que *pensa* que não fala castelhano, fique por isso liberto da pressão castelhanizadora sobre o seu próprio idiolecto; esta será menor, mas existirá e actuará pelos mesmos mecanismos que no falante não *aloglota* –basicamente psicológicos–, inserindo-se na percepção das relações de poder e na estratificação social, e na fuga à sensação de ridículo. Como não dispomos de metodologia que nos permita entrar em estudos deste tipo, e menos para épocas passadas, é difícil deslindar a mudança linguística de mecanismo interno da induzida externamente: na primeira são as próprias características da estrutura da língua e as razões de *economia* que produzem as mudanças; na segunda são factores de ordem psicológica, também com duplo mecanismo: a *imitação* (a iniciativa da mudança é do sujeito utente) e a *pressão* (o sujeito é coagido por outro); mas considerar externa a mudança linguística produzida por mecanismos psicológicos só é possível desde o pré-conceito filológico de que é *autóctone* e que é *externo*, que pode não existir na consciência do sujeito; não esqueçamos que a aprendizagem secundária da própria língua materna no sistema escolar produz-se pelos mesmos mecanismos de *imitação* e de *pressão*.

Parece que uma das estratégias do sujeito mergulhado numa situação de contacto de línguas, quer dizer, de comunidades humanas com diferentes modelos de relacionamento social, é a *hibridação*, a formação de uma modalidade linguística intermédia que represente um compromisso entre os dois pólos de influência que interferem no seu campo de acção¹⁰. Nesse comportamento costuma haver também uma componente lúdica, um exercício simultâneo de *cortesia* e de *ironia* consistente em ceder à pressão distanciando-se psicologicamente de quem a exercita. A hibridação abre ainda mais, como é óbvio, o campo de variabilidade dos sistemas linguísticos defrontados e cria um espaço neutro, permeável, favorecendo a mudança de língua dos sujeitos dispostos a dar esse passo.

Para tentarmos reconstruir as circunstâncias concretas em que ocorreu na sociedade medieval da Galiza o contacto entre o galego-português e o castelhano não podemos esquecer que a língua tem na nossa época o mesmo papel que a religião na sociedade medieval: é um elemento basilar na legitimação e no funcionamento da instituição estatal. Esta premissa obriga-nos a reconstruir o que pode ter sido a sensibilidade e a percepção medieval, distinta, em muitos aspectos, da dos dias de hoje. Deve ter havido uma diferença radical num ponto concreto: na Idade Média não parece ter havido rejeição pela mistura de códigos linguísticos, o que é um dos traços mais característicos e constantes do comportamento dos sujeitos nas sociedades desenvolvidas actuais.

O estudo da *penetração* do castelhano na Galiza revela-se problemático por ser um verdadeiro *campo de batalha* ideológico num aspecto decisivo da história do galego-português, sobretudo porque o que está em causa é o papel que pode ter desempenhado na –pretensa– ruptura da unidade linguística galego-portuguesa. Não esqueçamos que nesse ponto há basicamente duas posições: uma que considera que essa unidade nunca se rompeu totalmente, e aplica a denominação de galego-português à língua *própria* –actual– da Galiza, e outra que pensa que há uma época de língua unitária galego-portuguesa –até finais do séc. XIV– e que desde esse momento já é preciso reconhecer a existência de duas línguas diferentes. O postulado da independência linguística do galego assenta na interpretação de que o seu

¹⁰ Revela-se interessante, neste sentido, o estudo do castelhano escrito na Galiza em finais do século XV, e especialmente a *acomodação* de termos galego-portugueses que os autores consideram desnecessário substituir pelos correspondentes castelhanos; mesmo faz sentido comparar esse tipo de discurso com o de um galego actual castelhano-utente de língua materna galego-portuguesa: na fala deste último costuma manter-se inalterada uma parte do vocabulário vinculada às vivências e estados emotivos, como são os sufixos indicativos de carinho ou ternura, etc. No castelhano dos documentos notariais da segunda metade do séc. XV mantém-se a forma galego-portuguesa no vocabulário que designa realidades autóctones: *eira*, *malhar*, *tega*, etc. No castelhano de Vasco de Aponte, linhagista corunhês da primeira metade do séc. XVI, observa-se que o vocabulário de alguns campos semânticos, como o das relações de parentesco, mantém invariavelmente as formas galego-portuguesas.

relativo distanciamento dos falares portugueses setentrionais responde a factos de evolução interna do próprio sistema linguístico¹¹.

Para avaliar os elementos *objectivamente* castelhanos ou híbridos que se detectam nos documentos medievais da Galiza redigidos na língua autóctone surge sempre o problema de interpretar o tipo de relacionamento entre os dois romances, especialmente no que diz respeito à percepção do *status* relativo de um e outro na consciência dos sujeitos. Nesse campo concede-se certa importância aos topónimos: sendo na sua maior parte –especialmente os que se referem a entidades pouco importantes e a nomes de herdades– formas *unilingues*, por carecerem de equivalente tradicional vivo na língua *forânea*, pensa-se que a sua *tradução* segundo o padrão do vocabulário comum patenteia um processo de *dialectalização*: os utentes percebem o diferente *status* dos sistemas linguísticos em contacto e modificam as características distintivas do *inferior* para acomodá-las ao carácter solene e institucional do discurso, marcado pelo uso da língua *superior*. Os topónimos são um tipo de vocabulário que em princípio não deveria ver-se afectado: sendo vazios de significado conceptual, ao designar entidades referenciais únicas, não deveriam ser *traduzidos* se carecem, como na maioria de casos, de equivalente genuíno e reconhecível na língua *alheia*. É aí que radica o seu carácter *significativo*: a tradução constitui uma espécie de *hipercorreção*, um acto falhado, um passo em falso que revela a *intencionalidade* do locutor, descobre o preconceito que funciona na sua mente: a pretensão de solenizar o discurso por um procedimento de tradução, portanto o *status* superior que adquiriu a língua *alvo* na sua *percepção*. É evidente que as *castelhanizações*, no caso do galego-português da Galiza, têm diferente valor informativo segundo o tipo de texto em que se inserem: a tradução de topónimos num documento redigido em castelhano não constitui um dado significativo, pois podem ser formas criadas *ad hoc* no momento da redacção, talvez mesmo de maneira mecânica, por um locutor autóctone que traduz literalmente para a língua *institucional*. Mas um castelhanismo toponímico num documento redigido em galego-português pode considerar-se *significativo*: revela que as formas *traduzidas*, sentidas como mais solenes, já invadiram a oralidade do galego-português sendo percebidas pelos locutores como formas *próprias*. Esta classe de dados estaria a revelar, certamente, o pleno andamento de um processo de *dialectalização*¹².

¹¹ A corrente que postula a manutenção da unidade do galego-português até ao momento actual assume, como é óbvio, que tudo aquilo que nos falares galegos actuais é devido à interferência do castelhano não pode ser tomado em consideração para avaliar a distância relativamente aos falares portugueses setentrionais, porque nunca foi reconhecida a legitimidade do castelhano como elemento interveniente na *formação* do galego-português. Esta tendência desautoriza a atitude de considerar *galego* tudo aquilo que os utentes pensam que *é* e usam com pretensão de que seja; considera que a *fronteira* [normativa] entre galego-português e castelhano deve ser definida com critérios objectivos, não por sublimação da *percepção de fronteira* que tem o utente ingénuo.

¹² A conjugação dos dois factores (elemento castelhano + revitalização da variação autóctone) foi salientada por J. L. Rodríguez (1996:190): “A diferença entre uns e outros [textos medievais galegos e portugueses] radica na maior homogeneidade dos textos galego-portugueses elaborados em Portugal. Na Galiza, a variabilidade explica-se por *dialectalismo* e *castelhanização*”.

Os dados que deixam entrever o *status* atribuído a cada língua na percepção dos utentes são mais abundantes em épocas mais tardias, especialmente na segunda metade do séc. XV. Os documentos notariais revelam que o latim se manteve num plano de maior *solenidade* que o romanço durante toda a Idade Média, sobretudo porque surge nas partes mais formulárias e rituais –as que autorizam e autenticam o documento; na segunda metade do séc. XV temos indícios de que o mesmo tipo de contraste que se dá entre romanço e latim é transferido para a equação galego-português/castelhano; constata-se pelo aparecimento de assinaturas em castelhano de sujeitos identificados no texto em língua galego-portuguesa:

“E eu **Fernando de Çedeyra**, escriuano e notario suso dito, a todo o sobre dito en un con o dito Martín Afonso, notario, e con os ditos testigos, presente foy, esta escriptura con mao allea fis escripuir [...] **Fernando de Çedera** (signo)” (documento de 1481, A.H.P.O.–Caixa 9.831. Clero/Cela-Nova. Da edição de Ferro Couselo 1967: 155-179, nº 92).

“**Gómez Ares de Ferera**, notario”; no mesmo documento: “mosteyro de San Salvador de Ferrera [...] enno mosteiro de **Ferreira**...” (1482 Agosto 21; A.F.P., 3/192. Da edição de Fernández de Viana y Veites 1994: 253-254, nº 201).

A toponímia revela-se resistente à castelhanização, inclusive em documentos redigidos em castelhano na segunda metade do séc. XV:

“Don Fernando e doña Isabel por la gracia de Dios Rey e Reyna de Castilla, de Leon, de Aragon, de Sicilia, ...[...] Sepades que por parte del Abad, prior e monjes y conbento del monesterio de San Clodio, que hes en la diocesis de Orense, nos fue fecha relacion por su peticion que ante Nos en el nuestro Consexo presentaron, diciendo que don Bernaldino Sarmiento, conde de Rivadavia contra justicia y razón los tiene entrados los cotos, felegresias de Lebosende e de san Lorenzo de la Pena, e el coto de **Figueiredo** con las montañas, e el coto de Bobeda, e la freguesia (roto) Payo e la de santa Maria de Villarnaz (roto) de **Veyro**, (roto), san Juan de Abruciños, e las granjas e bouzas de Lamas, e la de Costana e (roto) e de **Oteiro**, e a d' Airas, e la Torre, e Orega, e **Cazcalleira**, e Montes e la granja de Reza...” (1486 Outubro 9; A. H. P. O., caixa 10.144. Da edição de Lucas Álvarez/Lucas Domínguez 1996: 710-712, nº 602).

Muito menos resistentes são as formas antroponímicas, inclusive as derivadas de topónimos:

“Anno del nasçimiento de noso Sennor Yhesu Christo de mille e quatroçentos e nobenta annos, **postromeiro** dia del mes de agosto. Don frey Rodrigo de Sangens, abade del monesterio de san Cloyo del Ribeyro d' Abya [...] aforraron a **Juan Nuñes de Heras** e a sua muller Teresa Afonso y a tres vozes

despues dellos, conbyen a saber el lugar que esta syto en la freyguisia de **santa Ougea de Eyras**, que se chama o lugar do Curral [...] de todo lo que Deus der en las dichas heredades, e las voces que despoyos de vos byeren, pagaram foro de carto, como sobredicho he, o pan partido en la **eyra** por dyante de nos o de noso mayordomo...” (1490 Agosto 31; A. H. P. O., Livros 686, Livro I de foros, f. 46 r. Da edição de Lucas Álvarez/Lucas Domínguez 1996: 718-719, nº 610).

Nos finais do século XV e começos do XVI surgem com certa frequência adaptações de revelam a interferência dos romances orientais no galego-português da Galiza:

Afonso de Toves [=Touves; no mesmo documento, linha 2, “Touves”] (1499 Dezembro 1; A.H.P.O., Osseira, nº 191. Da edição de Azevedo Maia 1986: 177-178, nº 89).

Roy Maçeras [=Maçeiras] (1501 Setembro 19; A.H.U.S., F. Univ., Série Hª, pergaminhos, nº 344. Da edição de Justo Martín/Lucas Álvarez 1991: 518-519, nº 383).

Martin de Fageras [=Fageiras] (1506 Fevereiro 27; A.H.U.S., F. Univ., Série Hª, pergaminhos, nº 347. Da edição de Justo Martín/Lucas Álvarez 1991: 525-527, nº 386).

Pero Ogia [=Ougia] (1532 Abril 17; A.H.U.S., F. Univ., Série Hª, pergaminhos, nº 368. Da edição de Justo Martín/Lucas Álvarez 1991: 574-576, nº 408).

A principal dificuldade que se coloca para a interpretação filológica da presença de formas castelhanas –quanto à génese– nos documentos da Galiza redigidos em galego-português é a sua interpretação. O conceito de *castelhanismo* é demasiado ambíguo, por polivalente. Em princípio teríamos de distinguir *formas castelhanas*, *castelhanizações* e *castelhanismos*. As primeiras corresponderiam ao uso da forma genuína do romance de Castela com consciência de ser tal: num discurso intencionalmente castelhano; as outras seriam o resultado da acomodação em castelhano de formas lexicais galego-portuguesas.

Castelhanização será a criação *ad hoc* de uma pretensa forma castelhana por adaptação fonológica ou morfológica a partir da original galego-portuguesa para ser inserida no discurso institucional em castelhano; essas adaptações podem ser puramente fantásticas, quer dizer, inventadas mecanicamente no momento de redigir o texto sem qualquer tipo de vigência na oralidade. A castelhanização típica é a *tradução* para castelhano de uma forma (toponímica, por exemplo) que carece de equivalente nessa língua, conhecendo-se exclusivamente a galego-portuguesa original: por exemplo *Seijo*, tradução de *Seixo*, topónimo ou apelido. Mas este con-

ceito é problemático para aplicá-lo às formas híbridas que aparecem muitas vezes em documentos em galego-português, nos quais, pela sua própria natureza, o esperável seria o processo inverso: acomodações em galego-português de formas originariamente castelhanas.

O aparecimento de uma forma toponímica *Orense* num documento não teria de ser um castelhanismo porque essa era a forma castelhana e pode verosimilmente ser usada como tal, inclusive num documento redigido em galego-português¹³. O problema é definir o estatuto das formas de origem *alheia* inseridas no discurso em língua *própria*; o critério que tende a aplicar-se é o da consciência do locutor: os vocábulos deixam de ser *alheios* (castelhanos, por exemplo) quando o emissor os sente e usa como *próprios* (castelhanismos). Um castelhanismo toponímico será aquele em que a forma castelhana ou híbrida substituiu a autóctone no discurso em língua *própria* e adquiriu na consciência dos utentes o *status* de forma *própria*¹⁴. Isto torna o problema insolúvel: não se pode saber que tipo de percepção tinham das variantes do *falar romance* nem o seu grau de indiferença perante a questão. A observação dos usos actuais na Galiza revela que quanto mais instruído for o falante maior é a sua preocupação por distinguir códigos linguísticos¹⁵: o utente espontâneo diz *carretera* (=estrada) sem detectar a monotongação de um ditongo *ei* que revela, na sensibilidade do utente-filólogo, a presença do castelhanismo; do mesmo modo usa *gasoliner*a no discurso em língua *própria*, mas é *gasolineira* a forma preferida pelo utente *culto*, reconstruída, evientemente, a partir da anterior. *Lo siento*, *buenos días*, *bueno*, *ojalá* são fórmulas que o utente mais espontâneo diz naturalmente sem mudar de código, mas um locutor instruído usá-las-á apenas no discurso em castelhano, e tradu-las por *Sinto-o*, *bôs dias*, *bueno* (aqui não!), *oxalá* ou *ogalhá* (!) no de língua *própria*. A indiferença do locutor *ingénuo* perante os pro-

¹³ O normal é que só os nomes de localidades importantes tenham dupla forma toponímica (*Mellid* e *Melide* ou *Milide* convivem desde a época medieval), mas encontram-se alguns castelhanismos toponímicos bastante surpreendentes: “o noso casar do **Campillo** [...] segundo que o el gaano de Mayor do Campello” (1396 Julho 9; A.F.P., 1/75. Da edição de Fernández de Viana y Vieites 1994: 103-104, nº 82); neste caso, aliás, surge o paradoxo de ter sido a forma antroponímica mais resistente à castelhanização do que o topónimo, quando normalmente ocorre à inversa. Note-se que nos documentos do pleito Tavera-Fonseca (1525-29) surge um sobrenome pessoal enigmático, *Campixo*, que cremos deve interpretar-se como castelhanização de *Campelo* por meio da acomodação com outro sufixo diminutivo, -*ijo*: *Pedro de Campixo* ou *Pedro Campixo*, vizinho da vila de Noia, vassalo do arcebispo de Santiago, cf. Rodríguez González 1984/II: 284 e 551.

¹⁴ As pessoas estão conscientes de que há *coisas* que se dizem em latim—ou em francês, inglês ou outra língua—que não faz falta ou não é costume traduzir, e isso não altera o carácter *genuíno* do discurso: *ipso facto*, *ex professo*, *de facto*; mas quando uma destas locuções se fixa na língua *própria* só o locutor culto mantém a consciência do seu carácter *alheio* quanto à génese: o *ingénuo* interpreta-as como locuções *normais*, idênticas a quaisquer outras, sem colocar-se o problema da sua origem.

¹⁵ Este facto está na origem do que se denomina *hiperenxebrismo* (=hipercasticismo): a reconstrução inadequada da forma *própria* de vocábulos que não precisam de ser modificados porque são idênticos nas variantes genuínas dos dois códigos linguísticos.

blemas da interferência de códigos é absoluta; não se importa se *Lo siento se deve* ou *não deve* dizer: só sabe que *se diz*; não sabe se é ou não é castelhano, mas é *galego*, porque *se diz*, como faz de facto ele próprio. Isto ocorre no momento presente: é verosímil que tenha ocorrido de forma muito mais radical na Idade Média. A percepção do locutor sobre o que faz parte da língua própria e o que é castelhano está condicionada pelo uso, mas não é filologicamente fiável, como é óbvio; para o locutor ingénuo **mai/nai** é língua *própria*, **madre** castelhano; em certas áreas geográficas **catro** é língua *própria*, **quatro** castelhano; **tolo**, língua própria, **louco**, forma inventada (!) –a partir do castelhano– pelo neo-falante que ignora a forma correcta da língua *própria*; na consciência do utente *espontâneo* **lindo** é palavra castelhana, como **repolho** [repolo] ou **andrajo** [farrapo] mas não são **silla** [cadeira], **malo** [mau] ou **carretera** [estrada]. Mas a percepção varia por áreas geográficas, pois o que nalgumas é **queijo**, ou **martelo**, noutras é já *queso*, *martillo*, formas usadas como próprias no discurso galego-português.

Devemos ter presente o duplo conceito de *castelhanismo*: o que filologicamente ou geneticamente é de origem castelhana –critério objectivo– e o que na consciência dos utentes é castelhano –critério subjectivo. Há muitos vocábulos que patenteam a necessidade desta distinção, como por exemplo a locução *el-rei*, fórmula de origem controversa, mas que os sujeitos galegos de língua galego-portuguesa terão percebido como *castelhana*, acomodando-se a *consciência linguística* a um uso que favorecia este tipo de consideração: o castelhano como língua da autoridade real, e depois língua solene, simplesmente; fica aberto o caminho para uma distribuição *funcional* dos códigos. O papel simbólico do rei é importantíssimo, na percepção popular, sobretudo como *fonte de justiça* –as Irmandades invocam a fidelidade ao rei para legitimar a sua rebeldia contra os abusos dos nobres.

O processo de dialectalização terá tido desde o seu início duas manifestações principais que podem servir-nos para, desde o momento presente, detectá-lo através das fontes historiográficas dos períodos históricos em que começou a produzir-se: a especialização das línguas em contacto para diferentes funções e a anulação do enriquecimento vocabular da língua subordinada, que deixa de ser independente e começa a adoptar de forma sistemática as inovações lexicais da língua superior. No caso dos topónimos autóctones que foram substituídos por castelhanismos toponímicos –o mesmo que imaginamos ter ocorrido com a locução *el-rei*– terá sido a especialização funcional dos códigos linguísticos a determinar a instalação do castelhanismo –subjectivo ou objectivo– na língua própria; o carácter unívoco e formal do acto jurídico terá motivado a exclusão de um dos sistemas e a apropriação completa das funções institucionais pelo outro; o rito formal e solene que confere vigência e validade a um documento é a *assinatura*, e constatamos de facto que esta se faz em latim até um dado momento, e já depois em castelhano. A referência ao rei como *el-rei* terá sido também de carácter ritual, formal e solene: independentemente de que a fórmula tenha tido origem castelhana ou leonesa¹⁶, na mente dos

utentes terá reforçado a percepção de que a língua adequada para designar o monarca era o castelhano, o que terá favorecido a sua especialização para os usos institucionais. No caso dos castelhanismos toponímicos é mais visível o percurso de ida e volta do processo: a forma galego-portuguesa é *traduzida* para castelhano para maior validade dos actos jurídicos em que é invocada, e depois *regressa* à língua própria como *castelhanismo*, ao ser usada como se fosse própria, provocando a queda em desuso do topónimo original¹⁷.

Não parece fácil encontrar provas irrefutáveis da castelhanização de topónimos galegos na época medieval, sobretudo porque, evidentemente, o facto de alguns deles aparecerem com formas castelhanas não significa que tivessem sido castelhanizados, mas que o autor ou redactor do documento usou a forma correspondente castelhana, viva presumivelmente no uso de Castela, sobretudo quando se trata de topónimos importantes, como são os nomes de cidades ou do próprio reino da Galiza. Do mesmo modo que *Cuenca*, *Sahagún* ou *Belorado* tinham *Conca*, *Sam Fagundo* ou *Belfurado* como formas galego-portuguesas reconhecidas e usadas, não nos pode parecer estranho que formas como *Orense* ou *Gallizia* tenham tido vitalidade, também na própria Galiza, desde o momento em que começa a haver grupos humanos que usam o castelhano como língua veicular habitual.

Parece claro que desde o século XIII Castela representa para o estamento nobiliárquico galego o centro de poder, e encontram-se muitas referências –sobretudo em testamentos– à ida para Castela obedecendo a convocatória do rei, para obrigações militares principalmente:

eu Pedro Fernandez Marques do mia erdade que hei uila Foz de Chamoso, en Franciam e en Castelon e en Galaens, en comenda e en guarda al comendador domino Suero Diaz de Loio; e sepades que **eu me uoo para Castela**; e se eu ala finir mando a la Orden de la Caualeria de Santiago por mia alma quanto ubier de toto parente e de todo home de mundo... (1241 Agosto 20; A.H.N., Sam Marcos, pasta 390, nº 199. Da edição de Novo Cazón 1986: 218-219, nº 20).

ego domnus Pelagius Fernandi, \miles/ de Rotario, **quando ivi ad dominum regem per mandatum et vocationem ipsius**, in primis mando corpus meum at animam meam monasterio sancte Maria Ursaria... ([1255] Outubro 2; A.H.N., pasta 1526, nº 20. Da edição de M. Romaní 1989/I: 704, nº 742).

¹⁶ Cf. a este respeito a informação que oferece Azevedo Maia (1986:646).

¹⁷ O que há de comum aos dois factos que apontamos como indícios de dialectalização é o aparecimento de *lacunas* vocabulares na língua própria que renuncia a ter formas divergentes das da língua institucional em determinadas áreas vocabulares; os termos da língua institucional passam a ser também próprios, por um processo de unificação selectiva dos dois códigos. O processo de castelhanização da língua autóctone deu-se com diferente ritmo e intensidade nas diversas áreas geográficas, e parece claro que o estudo do vocabulário comum pode esclarecer quais foram os focos irradiadores.

ego Arias Petri de Tabulata, **volens ire ad terram de Castella et ad curiam domini Regis**, mandam et testamentum meum facio in hunc modum (1257 Setembro 3; A.C.O., Monásticos, nº 528. Da edição de M. Romaní, 1989/II: 753-754, nº 790).

eu Pedro Eanes, escudeyro de Requeyxo, fillio de Johan Arias et de Orraca Oarez, **querendome yr para Castella buscar consello**, et temendome da morte e cobiiçando fazer prol et saude da mia alma, faço mia manda en esta maneyra, conven a saber: que se eu desta carreyra moyro ante que aca torne mando por mia alma [...] quanto ey et devo a aver [...] en terra de Asma et de Lemos et alur u quer que possam seer achados... (1279 Agosto 21; A.C.O., Monásticos, nº 1.314. Da edição de M. Romaní 1989/II: 1085-1086, nº 1141).

Sabeam todos commo eu, Gomez Perez de Guyllar scudeyro, sao et con mia saude et con todo meu entendemento qual me Deus dou, **entendendo yr a Castella** faço meu testamento et ordino de meus bees commo depoyos dia morte remeescam ordinadamente... (1316 Junho 28; A.H.N., pasta 1541, nº 15. Da edição de Romaní/Portela/Rodríguez/Vázquez 1993: 33-34, nº 1402).

Nos finais do século XV muitos nobres são obrigados a residir na Corte, o que parece ter sido um castigo e também um meio preventivo para evitar tanto os conflitos violentos entre eles como o abuso e rapina constante contra os mosteiros e os populares; Vasco de Aponte narra assim um episódio que diz ter ocorrido quando os Reis Católicos vieram à Galiza (1986: 201):

“Según las informaciones de Diego López de Haro, los reyes mandaron al señor Diego de Andrade y al conde don Sancho que se fuesen tras ellos a Castilla, y al conde de Altamira, por ser grueso y pesado, de piedad le mandaron quedar. Mas sucedió que un día acabando él de comer llegó un abad o prior, y no sé que le demandó; aborreciose y djóle: “Frade: irse an os hóspedes e comeremos lo galo”. Fuese el frayle a quejar al rey, y luego le mandaron que se fuese a Castilla dentro de tanto término so pena de muerte. Y así lo hiço”.

Parece necessário valorizar alguns factores que terão dificultado a substituição do galego-português pelo castelhano na elaboração dos documentos notariais na altura em que esta alternativa começa a ser praticável: (i) o peso de uma tradição de vários séculos, e (ii), ser o galego-português a língua da oralidade e a mais inteligível para as partes intervenientes nos actos jurídicos. Todavia, na segunda metade do século XV alguns documentos fornecem indícios de que uma das partes se identifica com (e pelo) castelhano enquanto a outra parece ter o galego-português como código de preferência:

Et **nos la abadesa** et pryora et monjas et conbento asy volo outorgamos **en las condycyones sobredy-tas**, et oblygamos a nos et aos veens do noso mosteiro a vos lo fazer de saoo et de pa_ et bos defender con el a dereito. Et non venderedes nen supynoraredes sen prymeyramente requeryrdes a nos et a noso mosteiro. Et **eu Alvaro da Quaral** asy o recebo et oblygo a myn et a meus veens ao con pryor et pagar **con as condyçós sobreditas** para min et para minas vozes (1482 Agosto 21; A.F.P., 3/192. Da edición de Fernández de Viana y Vieites 1994: 253-254, nº 201).

Nessa época surgen também indícios de que o castelhano é já percebido como marca de autoridade, o que terá sido uma consequência directa do seu lugar de preferência no uso institucional; é revelador que o artigo definido que nos primeiros momentos dá solenidade à evocação dos monarcas, *el-rei*, se tenha tranferido para a prática totalidade dos títulos que assinalam a posse de poder efectivo, tanto no âmbito civil como eclesiástico:

Sabean quantos esta carta de foro vyren commo nos, donna Marya López, abadesa do mosteyro de San Salvador de Ferera, con **la pryora** et monjas et convento, todas juntas en noso quabydoo con quanpaa tangyda... (1482 Agosto 21; A.F.P., 3/192. Da edición de Fernández de Viana y Vieites 1994: 253-254, nº 201).

Sepan quantos esta carta de foro viren como nos, donna María Lopeσ, abadesa do mosteiro de San Salvador de Ferreira, et **la priora**, mon-/jas et conbento do dito mosteyro (1483 Dezembro 8; A.F.P., 3/193. Da edición de Fernández de Viana y Vieites 1994: 254-256, nº 202).

Sepan quantos esta carta de foro viren como nos, dona María López, abadesa do moosteyro de San Salvador de Ferreira, et **la priora**, monjas et conben-/to do dito moosteyro... (1484 Janeiro 17; A.F.P., 3/194. Da edición de Fernández de Viana y Vieites 1994: 256-257, nº 203).

Sepan quantos esta carta de foro viren commo nos, dona Maria Loppeσ, abadesa do mosteiro de San Salvador de Ferrera, et **la priora** et convento do dito mosteiro... (1484 Março 2; A.F.P., 3/195. Da edición de Fernández de Viana y Vieites 1994: 258-260, nº 204).

Veja-se um documento em que se constata o alargamento do uso de *el* antecedendo o título como forma ritual de solenidade; o interesse radica na reprodução fidedigna do discurso *oral* reproduzido:

Queixa que dou gonçalvo de castro soy **al obispo** de diego ougia escudeiro.

Anno de mill e quatro çentos e oytenta e quatro annos a vynte e seys dias de junio ennos paaços **del corral** ante **el sennor obispo** gonçalvo de castro soy morador en penedo dou queyxa ao dito sennor obispo de diego ougea escudeiro disendo que este martes pasado o dito diego ougea chegara a el a huna

arada e lle disera que non morase mays enno dito lugar nin arase as herdades del senon que juraba que lle cortaria as maaos e os pees e que el disera que mantebese deus **al obispo** e que o dito diego ougea lle disera **quel obispo** non lle seria boo nin outro ningun que **bibise** enno obispado dourense e que **aun quel obispo** posee a outros caseyros ennos ditos lugares que nunca deles **el obispo** gozaria nin lebaria deles renta ningun que soubese el e que o disese el asy ao dito obispo e que lle lebase o recabdo delo e que asy pasara e outras asaz de palabras de que **el obispo** sacou juramento sobre un synal de cruz etç. Sy era asy verdade e el jurou que sy. Tests. **El provisor** pero guiral joan de deça arçediano de bubal afonso de salinas escudeiro fernando de tribino criado do dito sennor provisor (1484 Junho 26; A.H.P.O., sign. 243, *Registro de min juan garçia notário do conçello da çidade dourense...*, fol. 156-156v Da edição de López Carreira 1993: 44, nº 110).

Há que ter em conta, porém, que já em textos galegos do século XIII o artigo *el* de solenidade é utilizado para referir entidades diferentes do monarca; aparece, por exemplo, com “convento” num documento do mosteiro de Monte de Ramo de 1258: “a uos dō ffreyto Iohane abade de Môte de Rramo ea **el cōuēto** deffe lugar...” (A.H.P.O., Pergaminhos de Monte de Ramo, nº 3; editado por Ferro Couselo 1967/I: 19[23], nº 8).

Em documentos do século XIII e primeira metade do XIV encontram-se alguma formas que pressagiam já nessa altura algum tipo de *contacto linguístico* entre as duas línguas irmãs na Galiza:

τ ffe alguē da mja ljnagē τ que do meu aya d’auer cōtra algua daf coufal fobredjtal uéer aya a maldigō de **Dyos** τ a mjna... (1281 Julho 8; A.H.N., pasta 543, nº 20. Da edição de Azevedo Maia 1986: 75-76, nº 26).

Sabiã quantos eſta carta virē commo **yo** Maria Pereç, hermáa de Clemente Yaneç, raçoeyro de Mendoñedo.... (1310 Maio 10; A.H.N., pasta 1.109, nº 9. Da edição de Azevedo Maia 1986: 89-90, nº 35).

duaſ figuras de **peçes** e o letereiro do ffeelo dizia (1302 Outubro 20; A.H.P.O., Santa Clara de Alhariz, nº 29. Da edição de Azevedo Maia 1986: 143-147, nº 61).

enerro, Orēffe (e também Ourēſe), ouer [linha 6], oue [três vezes, linhas 8, 27 e 31], erederof [linha 33], mas eredeeyros [linha 22] (1290 Janeiro 20; A.H.N., pasta 1487, nº 14. Da edição de Azevedo Maia 1986: 138-140, nº 58)¹⁸.

¹⁸ Neste documento surge **Ares d’Alcantara**, que parece ter tido uma relação intensa com o testador, que lhe manda “o meu perpono, o meu lorigon, a ma gorgeyra τ cen mr. da guerra en dineyros polo que defef oue τ o meu capelo do fferro”. Parece provável, portanto, que o testador tenha passado uma parte da vida, talvez um período longo, em Castela: o seu amigo Ares d’Alcantara terá sido castelhano, a julgar pelo sobrenome. A representação com <o> dos ditongos <ou> pode ser lapso do escrivão (que nota *oue*, *ouer* em vez de *ouue*, *ouuer*, onde é fácil omitir um de dois grafemas idênticos).

Et se alguen de nossa parte ou destraya veer a corrunper esta vendiçon aga a nossa maldizõ et a de **Dos** [sic] (1291 Novembro 4; A.C.O., Monásticos, nº 1560. Da edição de M. Romaní 1989/II: 1163, nº 1225).

aga a mina maldizõ τ **eadidios** [sic] (1290 Junho 4; A.C.O., Monásticos, nº 1536; transcrição nossa. Edição do documento completo: M. Romaní 1989/II: 1155, nº 1216).

A situação da segunda metade do século XV apresenta muito menos dúvidas. Um documento editado por Ferro Couselo (1967: 200-204, nº 108), de 1491 (A.H.P.O., Cela-Nova, pasta I), constitui um exemplo revelador de como se defrontam dois poderes que neste momento representam o conflito de duas tradições, encenado de maneira dramática num duplo sentido: porque o choque aí narrado num documento jurídico tem a vitalidade e força evocadora de uma peça teatral; e porque uma das partes em confronto, o poder senhorial autóctone –dos mosteiros e pequena nobreza– muito cedo ia abandonar a fidelidade à tradição do país no que diz respeito à língua. Aí em presença do escrivão real *Aluaro Fernandes Vorrajo* (a onomástica é reveladora), o procurador do mosteiro de Cela Nova defende, em galego-português, o senhorio do convento sobre as localidades da Pousa e Castrelo, e requer ao senhor de Monterrei, da família dos *Çũñiga*, para que derrube a força que mandou erguer numa dessas terras. *Françisco de Çũñiga*, em nome do pai, *Pedro de Çũñiga, mi señor*, responde, provocador e desafiante, em castelhano, dizendo que a força está no seu senhorio e que não será derrubada, invocando em seu favor a autoridade dos reis de Castela. Um dos seus acompanhantes é *Gonçalo Rodriguez de la Pasera* “escribano de libros de molde”; o texto tem o vigor de uma obra literária, só que o drama que se encena é a realidade.

Referências bibliográficas

- Aponte, Vasco de (1986): *Recuento de las casas antiguas del Reino de Galicia*. Introducción e edición crítica con notas por M. C. Díaz y Díaz, J. García Oro, D. Vilariño Pintos, M. V. Pardo Gómez, A. García Piñeiro e M. P. Oro Trigo (Equipo ‘Galicia hasta 1500’) (Santiago de Compostela: Xunta de Galicia, Consellería da Presidencia).
- Fernández de Viana y Vieites, J. I. (1994): *Colección diplomática del monasterio de Santa María de Ferreira de Pantón* (Lugo: Diputación Provincial).
- Ferreiro, M. (1999) [1995]: *Gramática histórica galega. I. Fonética e Morfosintaxe* (Santiago de Compostela: Laivento).
- Ferro Couselo, X. (1967): *A vida e a fala dos devanceiros. Escolma de documentos en galego dos séculos XIII ao XVI*, 2 vols. (Vigo: Galaxia).

- Justo Martín, M. X. / Lucas Alvarez, M. (1991): *Fontes documentais da Universidade de Santiago de Compostela. Pergameos da serie bens do Arquivo Histórico Universitario (Anos 1237-1537)*. Edición diplomática (Santiago de Compostela: Consello da Cultura Galega).
- López Carreira, A. (1993): *De Ourense medieval. Rexistro de Xoán García (1484)*. Anexo 17 do *Boletín Auriense*.
- Lucas Álvarez, M. (1986): “El notariado en Galicia hasta el año 1300 (una aproximación)” (separata), em *Notariado Público y documento privado: de los orígenes al siglo XIV, Actas del VII Congreso Internacional de Diplomática* (Valencia: Generalitat Valenciana, Consellería de Cultura, Educació i Ciència).
- Lucas Alvarez, M. / Lucas Domínguez, P. P. (1996): *El monasterio de San Clodio do Ribeiro en la Edad Media: Estudio y documentos* (A Coruña: Ed. do Castro).
- Maia, C. de Azevedo (1986): *História do Galego-Português. Estado lingüístico da Galiza e do Noroeste de Portugal desde o século XIII ao século XVI. (Com referência à situação do galego moderno)* (Coimbra: I.N.I.C.).
- Martins, A. M. (1994): *Clíticos na História do Português. Apêndice documental: Documentos notariais dos séculos XIII a XVI do Arquivo Nacional da Torre do Tombo* (Lisboa: Imprensa Nacional-Casa da Moeda).
- Novo Cazón, J. L. (1986): *El priorato santiaguista de Vilar de Donas en la Edad Media (1194-1500)* (A Coruña: Fundación ‘Pedro Barrié de la Maza Conde de Fenosa’).
- Portas, M. (1993): *Língua e sociedade na Galiza* (A Coruña: Bahía).
- Rodríguez, J. L. (1996): “Sobre algumas formas e locuções conjuntivas nos textos galegos medievais”, em *Actas do IV Congresso Internacional da Língua Galego-Portuguesa na Galiza. Em Homenagem a Ferdinand de Saussure*. Vigo, 28 de Outubro-I de Novembro de 1993: 189-201 (Ourense: Associação Galega da Língua).
- Rodríguez, J. L. (2000): “Para um perfil das formas de tratamento *vostede/vosté... você*”, em Rodríguez, J. L. (coord.): *Estudos dedicados ao Prof. Ricardo Carvalho Calero (Reunidos e editados por José Luís Rodriguez)*, vol. I: 847-883 (Santiago de Compostela: Universidade de Santiago de Compostela / Parlamento de Galicia).
- Rodríguez González, A. (1984): *Las fortalezas de la mitra compostelana y los irmandiños* (A Coruña: Fundación ‘Pedro Barrié de la Maza Conde de Fenosa’).
- Romaní, M. (1989): *Colección diplomática do mosteiro cisterciense de Santa María de Oseira (Ourense) 1025-1310* (2 vols.) (Santiago de Compostela: Tórculo).

Romaní, M. / Portela, J. M. / Rodríguez Suárez, M. P. / Vázquez Bertomeu, M. (1993):
*Colección diplomática do mosteiro cisterciense de Santa María de Oseira
(Ourense) 1310-1399* (Santiago de Compostela: Tórculo).

Sponer, M. (1934): "Documentos antiguos de Galicia", *Anuari de l'Oficina Romànica
de Lingüística i Literatura*, VII: 113-192.

Sobre a vogal temática dos verbos galegos*

Francisco Dubert García
Instituto da Lingua Galega

Resumo:

Neste traballo pretendo demostrar o carácter léxico da vogal temática dos verbos galegos. Para iso, primeiro tentarei mostrar que a vogal temática carece de natureza primariamente flexiva; segundo, que os usos flexivos da vogal temática son secundarios e non diferentes dos usos flexivos da raíz dos verbos, polo que non alteran o estatus léxico deste formante; terceiro, que tódolos comportamentos anómalos das vogais temáticas nos verbos regulares poden ser explicados a partir de diferentes tipos de procesos fonolóxicos. De feito, a vogal temática será considerada resultado dunha condición de boa formación fonolóxica dos temas verbais. Desta maneira, a vogal temática nin sequera é considerada un morfema ou un signo mínimo, senón un mero formante fonolóxico do significante dun signo.

Palabras chave:

Verbo. Vogal temática. Derivación. Flexión.

Abstract:

The aim of this paper is to demonstrate the lexical character of the thematic vowel of Galician verbs. With this aim, first I will try to show that the thematic vowel lacks a primarily inflectional nature; second, that the inflectional uses of the thematic vowel are secondary and not distinct from the inflectional uses of the root vowels, so that their lexical status remains unaffected; third, that the anomalous behavior of the thematic vowels in the regular verbs can be explained by different types of phonological processes. In fact, the thematic vowel will be considered the result of a phonological well-formedness condition of the verbal themes. In that sense, the thematic vowel cannot be considered a morpheme or a minimal sign, but a mere phonological formative of the sign's significant.

Key words:

Verb. Thematic vowel. Derivation. Inflection.

Thematic vowel. *In the older IE languages, a vowel which is always present after certain lexical stems and before any inflectional suffixes. Nouns, verbs and adjectives may exhibit such a thematic vowel, and each of these may be divided into several subclasses depending on which of the several thematic vowels is present, plus a further subclass in which no thematic vowel is present. (Trask 2000, s.v. thematic vowel)*

* Agradezo os comentarios con que María Álvarez de la Granja, Sonia Colina e Elisa Fernández Rei contribuíron a mellorar este traballo. Os erros que aquí se conteñan son, naturalmente, só responsabilidade miña.

0. Introducción

Existe na conxugación dos verbos galegos un segmento denominado vogal temática que presenta diferentes interpretacións segundo os autores da nosa tradición gramatical que se consulten. Por exemplo, Porto Dapena (1972) non identifica na conxugación verbal un compoñente ó que chame vogal temática. A propósito de alternancias coma as de *ama-r* ou *ama-ba* fronte a *am-o*, *am-ei* ou *am-ou*, fala da existencia de dous alomorfos do “leisema” (secuencia de fonemas que expresa o contido xeral da acción): *ama-* e *am-*. Desta maneira, a vogal en cuestión é parte do formante léxico dun verbo,

poisque, se comparamos verbos como *podar / poder*, *mentar / mentir* [...], veremos que á alternancia *a/e/i* corresponden significados diferentes non no prano gramatical, senón no léxico, e, polo tanto, tales vocáis hainas que interpretar como pertencentes ao leisema (Porto Dapena 1972: 25).

Por esta razón, as terminacións *-ei* ou *-ou* de *amei* e *amou* son para el expoñentes indescompoñibles da P1 Singular do Indicativo Pretérito e a P3 Singular do Indicativo Pretérito, respectivamente (1972: 26); e os verbos do galego teñen dous alomorfos no ‘leisema’: na C-I, raíz + Ø (*am-*) e raíz + a (*ama-*); na C-II, raíz + Ø (*com-*) e raíz + e, i (*come-* ou *comi-*); e na C-III, raíz + Ø (*part-*) e raíz + i (*parti-*). En todo caso, o que me interesa salienta é o carácter estritamente léxico que Porto Dapena lle adxudica ó segmento que outros autores denominan vogal temática.

Antón Santamarina (1974: 14) xulga este segmento como “signo de clase mórfica. Mediante ellas podemos adscribir una forma verbal determinada a una de las tres conjugaciones”. Seguen esta proposta Álvarez, Monteagudo e Regueira, para os que a vogal temática “é un morfema que clasifica os verbos en tres clases mórficas ou conxugacións” (1986: 310); ou Freixeiro Mato, quen afirma que “constitui o signo da clase mórfica, pois por ela adscribemos unha forma verbal a unha das tres conxugacións” (2000: 308). Polo tanto, para todos eles a vogal temática parece ser o significante de tres signos que posúen como significante cadanseu conxunto de alomorfos en distribución complementaria (a selección de cada conxunto depende da conxugación a que se adscribe o verbo) e que teñen como significado, precisamente, a adscripción dun verbo a un destes modos de conxugación.

Pola súa parte, Fernández Rei (1998: 591) non só sinala que a vogal temática “indica a clase mórfica ou conxugación”, senón que tamén introduce un matiz interesante ó afirmar que esta “ás veces pode ter valores secundarios de modo e tempo *e/ou* de número e persoa”.

Álvarez e Xove (2002: 242), pola contra, sinalan que os verbos “distribúense en tres clases conxugacionais, atendendo ás variacións formais (variantes de

expresión) das combinacións da VT cos MF [vogal temática, morfemas flexivos, F.D.], variantes que se presentan en distribución complementaria”. Estes autores non lle atribúen valor sígnico ou morfemático á vogal temática. Desta maneira, e como vimos antes, dentro de cada paradigma conxugacional cada vogal temática contén un conxunto de variantes que se poñen en estreita relación cun determinado alomorfo dun sufixo flexivo.

Como vemos, existen visións diferentes da vogal temática. Con todo, creo que as interpretacións son máis ou menos casables e poden dar orixe a unha imaxe unificada. Na concepción que aquí vou mostrar, coincido con Porto Dapena sobre o carácter léxico da vogal temática, pero considerareina un elemento segmentable e independente da raíz; aceptarei de Santamarina (1974) o valor sinalador de clase conxugacional, aínda que rexeitarei o seu carácter morfemático ou sígnico; finalmente, desenvolverei a idea de Fernández Rei (1998) de que a vogal temática pode expresar, secundariamente, contidos gramaticais, só que insistirei en que, primeiro, este comportamento non é exclusivo da vogal temática e, segundo, non nos di nada sobre a súa natureza primariamente flexiva.

Para defender estes postulados, primeiro mostrarei a miña concepción da estrutura da linguaxe (§ 1) e como considero que se estrutura a compoñente morfolóxica (§ 2); só despois, exporei a miña análise da vogal temática (§ 3), seguida dunhas conclusións (§ 4). Como resultará claro, só me ocuparei dos verbos regulares, pois os verbos irregulares merecerían un estudo á parte (aínda que sospeito que as ferramentas conceptuais que aquí empregarei serán igualmente válidas para os verbos irregulares).

1. As palabras coma interface

Segundo a definición clásica saussureana de *signo*, este elemento confórmase por medio da asociación entre unha imaxe fónica mental ou *significante* e un concepto ou *significado*. Esta definición, en tanto que programática, resulta simple, pois, a pouco que pensemos, é fácil comprobar que a complexidade da información ofrecida por un signo pode ser moita. Analicemos, por exemplo, unha forma coma *comeron*.

O seu significado conceptual é algo parecido a ‘introducir algo na boca, mastigalo e tragalo’; porén, sabemos que xeralmente hai alguén que ‘introduce algo na boca, que mastiga ese algo e que o traga’: *os nenos comeron*; tamén sabemos que podemos especificar claramente que cousa é ese algo que ‘se introduce, se mastiga e se traga’: *os nenos comeron un queixo*. O elemento do mundo que realiza a acción tórnase na expresión lingüística en *axente* de ‘introducir algo na boca, mastigalo e tragalo’; e o elemento do mundo que é comido tórnase na expresión lingüística no

seu *tema*. A estrutura que contén o significado conceptual COMER (=‘introducir algo na boca, mastigalo e tragalo’) e os papeis semánticos do tipo *axente* ou *tema* denomínase *marco predicativo*:

(1) [COMER (axente, tema)]

Debe, ademais, existir algún tipo de información que nos diga con que tipo de axentes ou temas se pode combinar COMER: por exemplo, é preciso que COMER teña un axente [+animado]. Toda esta información que acabo de mostrar é de orde semántica. Con todo, a información que ofrece *comeron* non remata aquí. Sabemos que *comeron* se refire á P3 Plural do Indicativo Pretérito de COMER. Adoita chamarse *propiedades morfosintácticas* a cada unha das particularidades que nos permiten opor *comeu* (Singular) a *comeron* (Plural), *comen* (Presente) a *comeron* (Pretérito) e *categorías morfosintácticas* a cada un dos paradigmas que as agrupan (tempo, modo, persoa...). Fóra dos valores que engadan estas propiedades, é evidente que *comeron* comparte o seu significado conceptual con *comín*, *comerei* e *comerás*... O feito de que poidamos asociar este tipo de propiedades morfosintácticas a COMER, e non a COMIDA, débese a que COMER está lingüisticamente formalizado como un verbo (V), i.e., existe unha relación de coindexación entre o significado conceptual COMER e a etiqueta sintáctica V:

(2) COMER_i V_i

A información sobre as propiedades morfosintácticas (P3, Plural, Indicativo, Pretérito) e a información sobre a categoría léxica (V) é, en principio, de natureza sintáctica; de feito, debemos entender que as propiedades P3, Plural, Indicativo, Pretérito se adxudican á etiqueta sintáctica V e non ó significado conceptual COMER:

(3) COMER_i [V_i [P3, Plural, Indicativo, Pretérito]]

Nalgunhas teorías da morfoloxía (véxase Matthews 1991) acostuma a designarse co termo *lexema* a unha unidade constituída polo significado conceptual cando está marcado cunha etiqueta sintáctica, neste caso, $\sqrt{\text{COMER}}$; e denomínase *palabra morfosintáctica* (Anderson 1992 chama a esta entidade *representación morfosintáctica*) ó lexema determinado por calquera dos conxuntos de propiedades morfosintácticas que lle corresponde á categoría léxica en cuestión, no caso presente, $\sqrt{\text{COMER P3, PLURAL, INDICATIVO, PRETÉRITO}}$. A diferenza entre esta fórmula e a que propuxen en (3) ten un fundamento teórico: para min, as propiedades morfosintácticas son determinacións da etiqueta sintáctica V e non determinacións do lexema $\sqrt{\text{COMER}}$.

En tanto que verbo, *comeron* posúe tamén un *marco de subcategorización*, i.e., indica con que tipo de categorías sintagmáticas debe agruparse para construír

unidades maiores. Neste caso, o verbo pode combinarse cunha frase nominal e formar con ela unha frase verbal (FV):

$$(4) \quad FV \rightarrow [V_i + FN]$$

Convén indicar esta información porque, por exemplo, é posible que o verbo forme a súa FV botando man dunha FN introducida por unha preposición: *os nenos [pensaron [nese conto]]* (neste caso, o verbo debe indicar cal é a preposición que se usa); ou pode ser que seleccione outra oración: *os nenos [quixeron [comer un queixo]]*. Como vemos, a información contida no *marco de subcategorización* tamén é sintáctica. Pola súa parte, a FV asóciase con outra FN, que funciona de suxeito, por medio dun principio común da sintaxe, de modo que se dea orixe a unha oración:

$$(5) \quad O \rightarrow [FN + FV]$$

Nótese tamén como cada un dos papeis semánticos do marco predicativo debe corresponderse cunha destas FN. Nunha oración coma *os nenos comeron un queixo* a FN suxeito está relacionada co papel temático *axente*, mentres que a FN complemento directo está relacionada co papel temático de *tema*:

$$(6) \quad [\text{COMER}_i(\text{axente}_a, \text{tema}_b)] \\ FN_a [[V_i [\text{P3, Plural, Indicativo, Pretérito}]] FN_b]$$

Finalmente, é preciso un mecanismo que nos diga cómo se pronuncia (3), i.e., $\text{COMER}_i [V_i [\text{P3, Plural, Indicativo, Pretérito}]]$. Este tipo de información prodúcese a compoñente fonolóxica da lingua e será o obxecto deste traballo. No entanto, debo ir adiantando que a compoñente fonolóxica tamén está lonxe de ser simple e de estar desconnectada da compoñente semántica ou da sintáctica. En primeiro lugar, existe un vínculo entre o significado conceptual COMER_i , a categoría léxica V_i e unha secuencia de fonemas /kome/:

$$(7) \quad /kome/_{i_1} \quad \text{COMER}_{i_2} \quad V_{i_3}$$

Isto significa que COMER_{i_2} , /kome/_{i_1} e V_{i_3} son unidades que pertencen ás diferentes compoñentes da lingua, das que cada unha xera as súas propias estruturas lingüísticas. En segundo lugar, estas compoñentes fonolóxica, semántica e sintáctica, están relacionadas entre si por medio de interfaces, de modo que estas relacións de interface tamén contribúen a crear ou a condicionar estruturas fonolóxicas, sintácticas ou semánticas. Necesitamos, pois, algún mecanismo que sirva para construír o significado do complexo que representei en (3) a partir de /kome/_{i_1}. A presentación deste mecanismo require a exposición dunha pequena reflexión que ocupará o seguinte apartado.

Do visto, despréndese que a propia noción de signo, enriquecida como acabo de facer aquí, non pode ser outra cousa có resultado dun fenómeno de interface denominado *entrada léxica*, i.e., un elemento lingüístico que resulta do cruzamento de información sintáctica, semántica e fonolóxica: “alén de información fonolóxica e semántica, a entrada léxica inclúe polo menos información gramatical do estilo da categoría léxica, algúns trazos de subcategorización e irregularidades morfolóxicas ou sintácticas” (Trask 1993, s.v. *lexical entry*). Neste sentido, Jackendoff (2002: 131) chega a afirmar que o propio lexicón debe ser visto como parte das compoñentes de interface.

2. Morfoloxía flexiva e morfoloxía derivativa

Na concepción que aquí se defenderá, a morfoloxía é unha compoñente da gramática resultado de relacións de interface que *no seo das entradas léxicas* se establecen entre as compoñentes semántica e sintáctica nun lado e a compoñente fonolóxica do outro. Noutras palabras, por morfoloxía entenderei as variacións na forma fonolóxica das entradas léxicas que ocorren parellas ás variacións detectadas na súa estrutura semántica ou sintáctica. A morfoloxía ocúpase das relacións formais existentes entre as diversas formas de palabra (sexa dentro da mesma entrada léxica, sexa entre entradas léxicas diferentes) e da estrutura interna das propias formas de palabra. É esta interrelación entre estas tres compoñentes o que define a morfoloxía¹. A morfoloxía adoita dividirse en *derivación*, *flexión* e *composición*. A seguir, comentarei lixeiramente as nocións de flexión (2.1), derivación (2.2) e os procesos morfolóxicos por medio de que se expresan estas dúas partes da gramática (2.3). Ignorarei a composición porque é unha cuestión sobre a que non é preciso reflexionar neste estudo.

2.1. A natureza sintáctica da flexión

Entendo a morfoloxía flexiva como o estudo das variacións na forma de palabra dunha entrada léxica asociadas a variacións no contexto sintáctico, i.e., a morfoloxía flexiva é a compoñente da gramática que nos mostra como a sintaxe afecta á forma da palabra. Para comprender plenamente esta afirmación fai falta insistir no carácter sintáctico das propiedades morfosintácticas. Este feito ten dúas implicacións que terei que defender a seguir: primeira, propiedades coma [P3,

¹ “Un campo de estudio coherente defínese máis pola a existencia dun conxunto de principios ben definido e unitario ca pola existencia dun único conxunto de obxectos ós que eses principios poderían referirse. Disto resulta, con todo, que se varias subpartes da gramática se sobrepoñen de maneiras significantes, entón as propiedades dos elementos compartidos serán só descritas en termos dunha complexa interacción de principios de distintas áreas; e a comprensión de calquera área só pode proceder conxuntamente coa nosa comprensión das outras” (Anderson 1982: 572).

Singular, Indicativo, Pretérito] non forman parte do *significado conceptual nin gramatical* de V; segunda, as operacións que implican propiedades morfosintácticas son, en principio, de índole sintáctica (nótese que me estou a referir ás propiedades, non ós elementos ou operacións con que se representan na fonoloxía).

Anderson (1988: 167-168; 1992: 82-83; en Stump 1998: 21-26 aparece unha clasificación similar) sinala que as propiedades morfosintácticas poden agruparse en catro categorías diferentes:

- a) Propiedades de configuración, que son asignadas en función da estrutura sintáctica superior en que unha palabra A se integra: en latín, un substantivo que ocupe a posición núcleo da FN suxeito dunha oración terá a propiedade Nominativo e un substantivo que ocupe a posición de núcleo da FN complemento directo dunha FV terá a propiedade Acusativo.
- b) Propiedades de concordancia, que se lle asignan a unha palabra A en referencia ás propiedades doutra palabra B situada dentro da mesma estrutura sintáctica: un adxectivo (A) que funcione de modificador dun substantivo (B) dentro dunha FN concordará en xénero e número.
- c) Propiedades de frase, que se asignan ós dominios das frases e que se realizan dentro dunha das palabras que aparecen nesa proxección: é o caso da aparición das propiedades Indicativo Pretérito no verbo de *comín queixo*, que permite inserir este verbo nunha oración principal (non hai oracións principais en galego que non presenten propiedades de modo e tempo).
- d) Propiedades inherentes, que son características predeterminadas nunha entrada léxica concreta e que deben ser accesibles ás regras de concordancia: o substantivo *cama* contén a propiedade Feminino e impón concordancia en xénero feminino ós adxectivos que o modifican, coma en *cama fea*, *cama feita*.

Segundo Anderson, mentres que as propiedades de configuración, de concordancia e de frase só poden ser asignadas por unha regra sintáctica, as propiedades inherentes deben ser vistas pola sintaxe, pois son disparadoras das regras sintácticas de concordancia. Estes dous feitos indican a natureza fundamentalmente sintáctica das propiedades morfosintácticas. Esta natureza é perfectamente clara na concordancia, que é vista decote como unha marca de relación sintáctica entre os elementos que conforman unha categoría sintagmática: a concordancia “é un tipo de covariación entre elementos relacionados sintagmaticamente. A concordancia é un indicativo de que elementos están relacionados e pode servirnos de indicio sobre a extensión dun constituínte” (García-Miguel e Cabeza 2000: 559). Martinet (1991: 119), pola súa parte, distingue os *monemas funcionais* (coma as preposicións e os

casos) dos *monemas non funcionais*; estes últimos divídense, á súa vez, en *modalidades* (coma o artigo ou a propiedade Plural nos substantivos ou a propiedade Copretérito nos verbos) e *monemas léxicos*. Os monemas funcionais (preposicións e casos) serven para sinalar as relacións que un sintagma ten con outro superior (polo que manifestan un *valor centrífugo*), mentres que as modalidades actúan dentro do sintagma precisando o significado dos *monemas léxicos* (polo que posúen un *valor centrípeto*). Como se ve, é unha diferenza similar (aínda que non equivalente) á que existe entre as propiedades de concordancia e as propiedades de configuración...

Agora ben, cómpre sinalar que as propiedades son tamén de natureza morfolóxica, pois a presenza de determinadas propiedades morfosintácticas determinando unha categoría léxica acostuma a estar correlacionada con diferentes tipos de estruturas fonolóxicas nas formas de palabra; e é esta correlación entre a sintaxe e a fonoloxía o que crea a relación de interface que dá orixe á morfoloxía flexiva.

Por outra parte, quero insistir en que as propiedades morfosintácticas non constitúen en por si parte do significado, do mesmo xeito que un suxeito non forma parte do significado dunha oración ou que a etiqueta verbo non implica un significado de acción. Así coma un suxeito pode ser a vía de expresión dun axente en *eu como*, dun tema en *o pan é comido* ou dun experimentador en *o leite ferve...*; verbo pode ser unha categoría léxica adscrita a un significado conceptual de acción, coma en *ir*, proceso coma en *quecer*, estado coma en *ser*; as propiedades morfosintácticas do verbo galego poden ir asociadas a diferentes significados gramaticais (xeralmente denominados *valores* na nosa tradición lingüística). Compárese se non:

- (8) a) *Serán as sete en punto cando veña*
 b) *Agora serán as sete en punto*
 c) *Dixo que volvería*
 d) *Aquel home sería máis vello ca min*
 e) *Axudaríache se puidese*

Como se pode ver, *serán* en (8a) significa futuro con respecto ó presente e certeza, *serán* en (8b) significa presente e incerteza, *volvería* en (8c) significa futuro con respecto a unha acción pasada e certeza, *sería* en (8d) significa incerteza nun punto do pasado e *axudaría* en (8e) significa simplemente irrealidade, pois o significado temporal non o ofrece a gramática, senón o contexto (véxase Santamarina 1974: 120).

O que estes feitos mostran é que as propiedades morfosintácticas non constitúen significados gramaticais, non son o significado gramatical do verbo nin do lexema verbal, senón que colaboran, xunto coa fonoloxía, na expoñencia de significados

gramaticais. Polo tanto, as propiedades morfosintácticas, elementos da mesma natureza sintáctica cá etiqueta V, manteñen tamén relacións de interface co mundo do significado. Unha representación da estrutura modo-temporal de *serán* en (8a) podería ser a seguinte:

- (9) V_a [Futuro]
SER_a, FUTURO, REAL
/seraN/_a

2.2. A natureza léxica da derivación

Mentres que os procesos flexivos producen realizacións diferentes das entradas léxicas, a derivación é un fenómeno que engloba a creación de novas entradas léxicas, de xeito que unha entrada léxica pode dar orixe a outra por medio de modificacións ben na súa categoría léxica, ben no seu marco predicativo, ben no seu marco de subcategorización, ben no seu significado conceptual, ben na súa forma de palabra, ben en todos e cada un dos compoñentes da entrada léxica, ben só nalgúns deles, ben nun só (Matthews 1991, Anderson 1992, Beard 1998). É preciso indicar que no galego, coma noutras linguas, é posible encontrarmos formas de palabra derivadas que presentan o mesmo significado conceptual, os mesmos marcos temáticos ou de subcategorización e a mesma categoría léxica cás súas bases, de modo que se producen cambios na forma de palabra que non van unidos a ningún outro cambio nas outras compoñentes da entrada léxica. Unha boa reia de exemplos encontrámola en González Seoane (1999: 484): *espapallar ~ espapar, esfurricallar ~ esfurricar, alcanñar ~ alcanzar*. Por esta razón, débese indicar que a derivación non implica necesariamente un cambio conceptual ou sintáctico: é abondo que haxa un cambio formal². Desta arte, en tanto que afecta a unha, a algunhas ou a tódalas compoñentes da entrada léxica, a derivación non só é un proceso de natureza estritamente léxica, senón que nos mostra que tipos de fenómenos deben ser considerados de natureza léxica.

Vexamos un exemplo. Unha parte das entradas léxicas de *comer* e *comedor* ('que lle gusta comer') e *comedoríño* poden ter unha forma máis ou menos similar á que ofrezco en (10):

- | | | | |
|---------|-----------------------------------|-------------------------|----------------------------|
| (10) a) | COMER _i (axente, tema) | b) COMEDOR _j | c) COMEDORIÑO _k |
| | V _i [+FN] | Adx _j | Adx _k |
| | /kome/ _i | /komedor/ _j | /komedorij/ _k |

² Cómpre sinalar que González Seoane mostra dúbidas de que os membros das parellas pertencen a diferentes variedades lingüísticas; tamén cumpriría explorar as repercusións teóricas de que isto así fose, quizais non tan problemáticas como podería pensarse a priori.

Como vemos, existen varias diferenzas. Comezarei por explorar primeiro as que existen entre (10a) e (10b). Primeiro, mentres que $COMER_i$ pode ser parafraseado como ‘introducir algo na boca, mastigalo e tragalo’, $COMEDOR_j$ glósase como ‘que lle gusta introducir algo na boca, mastigalo e tragalo’, de modo que conceptualmente se perde o valor de ‘acción’ e adquírese o de ‘característica’. Segundo, $COMER_i$ está coindexado cunha etiqueta sintáctica V_i e aparece nas posicións sintácticas propias de V , $COMEDOR_j$ está indexado coa etiqueta sintáctica Adx_j e aparece nas posicións sintácticas propias de Adx . Terceiro, como consecuencia do anterior, (10a) xera un conxunto de formas propias do seu paradigma: *comerei, comería, comera...*, mentres que (10b) contén outro conxunto de formas caracterizadas por unhas propiedades morfosintácticas distintas: *comedor, comedora, comedores, comedoras*. Cuarto, $COMER_i$ ten unha forma de palabra básica /kome/, mentres que $COMEDOR_j$ ten unha forma de palabra básica en parte distinta en parte similar: /komedor/. Pola súa parte, as únicas diferenzas entre (10b) e (10c) son a presenza do sufixo *-iñ-* e o valor semántico expresivo que gaña a forma derivada (compárese *o neno é comedor* vs. *o neno é comedoriño*).

Por estas razóns, a derivación é un proceso de natureza léxica. Como sinala Anderson (1992: 185):

Por unha parte, as descrições estruturais das regras flexivas refírense a propiedades da representación morfosintáctica, mentres que as descrições estruturais das regras derivacionais refírense ó contido das (clases de) ítems léxicos. Pola outra, os cambios estruturais das regras flexivas afectan só á súa forma fonolóxica, mentres que as regras derivacionais tipicamente tamén efectúan modificacións tanto na semántica coma na sintaxe dos ítems.

Isto significa que a flexión, i.e., a inserción das propiedades morfosintácticas pode producir cambios na forma fonolóxica da entrada léxica, mentres que a derivación crea unha nova entrada léxica a través de cambios na forma fonolóxica e/ou en calquera das súas outras compoñentes (categoría léxica, marco de subcategorización etc.).

2.3. Morfoloxía flexiva e morfoloxía derivativa

Ata agora estiven centrado en mostrar a natureza sintáctica da flexión e na natureza léxica da derivación. É agora o momento de reflexionar sobre as repercusións que estes cambios teñen no plano fónico. De acordo co visto ata agora, xa pode afirmar que será obxecto da *morfoloxía flexiva* o conxunto de cambios na estrutura fonolóxica que teñen xustificación na flexión e obxecto da *morfoloxía derivativa*

os conxuntos de cambios na estrutura fonolóxica que están relacionados coa derivación, tal e como derivación e flexión foron descritas arriba.

Como xa sinalai, dentro da entrada léxica *comer* a secuencia /kome/_i é a forma básica de palabra que se corresponde coa etiqueta V_i e co significado conceptual COMER_i, do mesmo xeito que na entrada léxica *comedor* a secuencia /komedor/_j é a forma básica de palabra que se corresponde co significado conceptual de COMEDOR_j e coa etiqueta Adx_j. É evidente que a entrada léxica *comedor* relaciónase coa entrada léxica *comer* a través dunha regra derivativa de formación de palabras. A derivación atangue variacións na compoñente sintáctica, semántica e fonolóxica. En virtude da categoría léxica que posúa cada unha das entradas léxicas, pode pasar a ser flexionada, i.e., pasa a formar parte dun *paradigma* no que se obteñen por flexión realizacións da entrada léxica. A flexión que recibe depende da lingua: en galego as preposicións, conxuncións ou adverbios nunca son flexionados, mentres que en bretón, por exemplo, as preposicións deben establecer concordancia cos pronomes persoais. Así, /komedores/ é a expresión do masculino plural da entrada léxica *comedor*, mentres que /komes/ é a expresión da segunda persoa de singular do presente indicativo de *comer*. Como vemos, unha cousa é a forma básica de palabra, /kome/_i e /komedor/_j, e outra é a forma de expresión das entradas léxicas plenamente flexionadas, /komes/ e /komedores/, respectivamente. No caso de *comes*, o proceso flexivo consistiu en implementar as propiedades [P2, Singular, Presente, Indicativo] a V_i; no caso de *comedores*, o proceso conduciuse por medio da implementación das propiedades [Masculino, Plural] a Adx_j. Como se sinalou arriba, a adxunción destes conxuntos de propiedades produciu tamén unha serie de cambios no significante, de forma que de /kome/_i debeu pasarse a /komes/ e de /komedor/_j a /komedores/. Disto despréndese que tamén desde un punto de vista morfolóxico /komes/ e /komedores/ son formas máis complexas ca /kome/_i e /komedor/_j. Reservarei, seguindo a terminoloxía de Matthews que xa presenteí, a etiqueta de *forma de palabra* para a secuencia de fonemas que serven para expresar unha entrada léxica efectivamente empregada nun contexto sintáctico, i.e., unha entrada léxica que ten a súa categoría léxica enriquecida coas diferentes propiedades morfosintácticas que sexa susceptible de portar.

Desta maneira, fíame falta un nome para a forma básica de palabra que aparece como significante na entrada léxica (no noso caso, /kome/_i e /komedor/_j). Chamareille *tema* á forma básica da entrada léxica: a secuencia de fonemas que está coindexada cun significado conceptual (ou gramatical) e cunha etiqueta sintáctica. Desta maneira, /kome/_i e /komedor/_j son os temas das entradas léxicas *comer* e *comedor*, respectivamente. A diferenza morfolóxica entre /kome/_i e /komedor/_j reside no seu grao de complexidade e é a morfoloxía derivativa a que debe dar conta dos cambios operados na forma fonolóxica. Nun caso coma o que estamos tratando, a complexidade formal vai asociada a unha complexidade

semántica; e do mesmo xeito que /kome/_i está incluído en /komedor/_j, COMER_i está incluído en COMEDOR_j (COMER_i ‘introducir algo na boca, mastigalo e tragalo’ vs. COMEDOR_j ‘que lle gusta introducir algo na boca, mastigalo e tragalo’). É por isto que podemos dicir que /komedor/_j é derivado de /kome/_i.

Chamareille *base* ós elementos sobre os que se pode construír un novo tema: desta maneira, /kome/_i é base de /komedor/_j e /komedor/_j é base de /komedorij/_k. Polo tanto, un tema é unha secuencia que pode recibir as marcas formais da flexión e unha base unha secuencia que pode recibir as marcas formais da derivación: evidentemente, /komedor/_j pode ser á vez *base* e *tema*. Con todo, esta clasificación non esgota tódalas posibles unidades, pois queda aínda un concepto que debe ser sinalado: pode existir unha parte da forma de palabra que permanece inmutable ó longo de tódolos procesos de derivación. Pensemos en *branco*, *branca*, *esbrancuxado*, *branqueas*...; a pesar de tódolos procesos de derivación e flexión, existe unha parte do significante que permanece inalterada, nomeadamente /braNk/. Esta é a *raíz*. Unha raíz pode ser base (se é susceptible de ser modificada nun proceso de derivación: *branquexar*) ou tema (se pasa por un proceso de flexión: *branc-o*, *branc-a*).

Do dito ata aquí, non fai falla repetir que por procesos morfolóxicos entendo procesos fonolóxicos motivados pola derivación e pola flexión. Existen diferentes tipos de procesos morfolóxicos, i.e., as linguas poden empregar diferentes tipos de operacións fonolóxicas para expresar a derivación e a flexión, pero cómpre desde xa sinalar que non é necesario que exista un conxunto concreto de operacións fonolóxicas destinadas á flexión e un tipo de operacións diferentes destinadas á derivación: tanto a flexión coma a derivación poden empregar un mesmo tipo de ferramentas fonolóxicas. Como sinala Stump:

Na lóxica da flexión, nada exclúe a posibilidade de que a flexión poida implicar o mesmo tipo de operacións formais cá derivación; máis ben, nada exclúe a posibilidade de que o mesmo tipo de operación poida servir para unha función derivativa nuns casos e para unha función de flexión noutros (1998: 19).

Polo tanto, flexión e derivación poden valerse do mesmo repertorio de procesos morfolóxicos: a afixación, a reduplicación, a apofonía etc. Por exemplo, a prefixación co formante *a-* é derivativa en galego en palabras coma *anormal*, *asocial*...; e é flexiva co formante *é* en grego en palabras coma *é-ly-es* ‘desatabas’, pretérito imperfecto de *lyo* ‘desatar’. A infixación é derivativa en verbos do chamorro coma *hasso* ‘pensar’ vs. *h-in-asso* ‘pensamento’, ou *faisen* ‘preguntar’ vs. *f-in-aisen* ‘pregunta’; e flexiva en latín en verbos coma a nasal de *ru-m-po* ‘rompo’ vs. *rupi* ‘rompín’ ou *vi-n-co* ‘venzo’ vs. *vici* ‘vencín’. Pode consultarse un completísimo estudo sobre os procesos morfolóxicos e sobre o tipo de formantes

que se usan para construír as formas de palabra no traballo de Pena (1991), de onde tomo algúns destes exemplos.

2.4. A falta de adecuación entre as propiedades morfosintácticas e os seus expoñentes

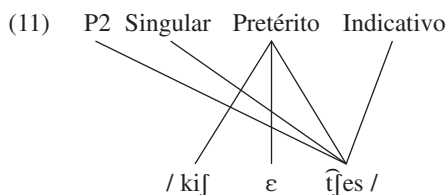
Antes de presentar a miña proposta de análise da vogal temática dos verbos galegos, quero sinalar dous feitos que se refiren ó valor sinalador dos formantes morfolóxicos. É común pensar que a morfoloxía consiste sobre todo na adición de *morfemas* a unha forma inicial máis simple co fin de construír unha palabra máis complexa. As series /kome/_i > /komedor/_j > /komedorij/_k, /kome/_i > /kome/N/_i ou /komedor/_j > /komedores/_j responden a este esquema ideal: nelas observamos unha engádegas no material significativo acompañada de modificacións quer no significado conceptual, quer na implementación de propiedades morfosintácticas; os morfemas non son outra cousa ca esas porcións de significativo asociadas a significados ou a propiedades morfosintácticas. Segundo se empreguen para a flexión ou para a derivación, podemos clasificar os morfemas coma flexivos, coma no caso de /komedor/_j > /komedores/_j, ou derivativos, coma no caso de /kome/_i > /komedor/_j > /komedorij/_k.

Con todo, é un fenómeno común a inadecuación entre o plano sintáctico, semántico e fonolóxico. Pensemos nos repetidos exemplos ingleses de conversión coma o de *doubt* ‘dubidar’ > *doubt* ‘dúbida’, onde existe un proceso de derivación que crea unha entrada léxica, *doubt* ‘dúbida’, a partir doutra, *doubt* ‘dubidar’, sen que, aparentemente, haxa ningún tipo de modificación do significativo (e polo tanto, ningún tipo de adición de material fonolóxico). E, no que á flexión se refire, pensemos no tema /kome/ da entrada léxica *comer*... Neste caso, vemos que, sen que aparentemente se engada ningún tipo de material fonolóxico, se emprega o simple tema da entrada léxica de *comer* para expresar a segunda persoa de singular do imperativo: /kome/; máis aínda, vemos que a terceira persoa de singular do presente de indicativo non se expresa coa adición de ningún material fonolóxico, senón coa modificación (neste caso, abaixamento) da vogal radical: /kɔme/. Deste último feito pretendo tirar dous corolarios.

³ Hai autores que nestes casos falan de *morfos cero*: no caso de /kome/, debería existir un morfo cero para a propiedade Imperativo e un morfo cero para P2 Singular; no caso de /kɔme/, un morfo cero para as propiedades Indicativo Presente e outro para P3 Singular; no caso do inglés, un morfo cero que transforma *doubt* ‘dubidar’ en *doubt* ‘dúbida’ (V > N), e outro morfo cero que transforma *group* ‘grupo’ en *group* ‘agrupar’ (N > V). Ergo, o morfo cero pode ser significativo de todo, incluso de propiedades morfosintácticas en oposición paradigmática. Exista ou non un morfo cero, o caso é que a terceira persoa do presente de indicativo de *comer*, /kɔme/, non se distingue da segunda persoa do imperativo, /kome/, pola adhesión de ningún tipo de material fonolóxico, senón por unha diferenza no grao de abertura da vogal radical; e o certo é que isto lle pasa á maioría dos verbos da segunda e da terceira conjugacións (Dubert 1999a para unha análise).

O primeiro, e para min máis importante, é que no caso de /kɔme/, a modificación da vogal radical ten unha función flexiva; a importancia deste feito radica precisamente niso: unha parte da forma de palabra de natureza léxica (pois a raíz é unha parte tradicionalmente relacionada coa expresión do significado conceptual) é modificada para expresar contidos *sintácticos*. Polo tanto, a morfoloxía flexiva, un conxunto de operacións fonolóxicas ligadas a condicionamentos sintácticos, pode operar sobre a mesma raíz, o elemento de natureza léxica prototípico nunha forma de palabra. Isto non é estraño se pensamos nos casos de supletismo coma *son, es, vou, imos* etc., en que toda a raíz serve para indicar tamén propiedades morfosintácticas⁴.

O segundo corolario divídese en dúas partes: a) a expresión dun mesmo contido *sintáctico* pode realizarse en diferentes compoñentes da forma de palabra; e b) unha mesma compoñente da forma de palabra pode expresar contidos sintácticos diferentes (isto é sistemático nos casos de morfemas *portmanteau*: en *cantase*, /se/ asóciase a Subxuntivo e a Pretérito). Conforme a isto, o grao de abertura da vogal radical sinala Indicativo e Presente en /kɔmes/ fronte a /komas/, en /kɔme/ (e) fronte a /koma/ e /kome/ (ti) e en /kɔmeN/ fronte a /komaN/; mentres que indica P2 Singular, P3 Singular e P3 Plural en *comes, come e comen* fronte a *como, comemos, comedes*. Por súa vez, P2 Singular está expresado tamén por /s/, P3 Singular por nada (ou por un morfo cero, se se quere empregar este recurso) e P3 Plural por /N/. Ademais, Indicativo e Presente sinálanse tamén pola presenza da vogal temática /e/: compárese /kɔmes/, /kɔme/, /komemos/, /komedes/, /kɔmeN/ con /komas/, /koma/, /komamos/, /komades/, /komaN/... Este fenómeno da falta de adecuación entre as propiedades morfosintácticas e os seus expoñentes pode verse representado en (11) coa forma do verbo irregular *quixeches*, P2 Singular do Indicativo Pretérito de *querer*:



Como vemos, /ʃes/ emprégase nas P2 do Singular do Indicativo Pretérito fronte a /s/, que se emprega no resto: *quixeches* fronte a *queres, queiras, quererías, quixeras...*; /e/ emprégase nas formas arrizotónicas no Indicativo Pretérito, no

⁴ Esta idea non é en absoluto nova na lingüística galega. Respecto destas alternancias vocálicas xa afirmou Porto Dapena que nos achamos “diante dun caso en que o sentido gramatical dáse xuntamente co lexicai, resultando imposible arredar na secuencia ou cadea fónica un do outro” (Porto Dapena 1972: 28).

Antepretérito e no Subxuntivo Pretérito: *quixemos, quixestes*, fronte ó /e/ de *queremos, queredes* no Indicativo Presente; por fin, a raíz /kiʃ/, que se emprega no Indicativo Pretérito, Antepretérito e Subxuntivo Pretérito. Polo tanto, a propiedade Pretérito está expresada por tres medios, /kiʃ/, /e/, e /tʃes/, de modo que unha mesma propiedade se relaciona con tres expoñentes (un deles, quizais unha raíz supletiva, véxase Dubert 1999a, claramente asociada á *parte* léxica da forma de palabra); mentres que o formante /tʃes/ se vencella a catro propiedades diferentes: P2, Singular, Pretérito e Indicativo.

3. O lugar da vogal temática dos verbos na lingua galega

De acordo co que se vén sinalando na nosa tradición gramatical sobre a morfoloxía verbal, a estrutura dun verbo galego debe conter unha raíz e unha vogal temática, conformando ambas o *tema* do verbo, seguidas dunha serie de sufixos expoñentes de propiedades morfosintácticas (véxase, por exemplo, Fernández Rei 1998). Aínda que dou por boa tal estruturación, quero salientar dous puntos. Primeiro, esta definición de *tema* non coincide exactamente coa que ofrecín arriba, pois un *tema verbal* pode estar formado por máis elementos cá raíz e a vogal temática: é o caso da entrada léxica de *enriquecer*, onde encontramos un tema formado por un prefixo /eN/, unha raíz /rik/, un sufixo /eθ/ e unha vogal temática /e/. Segundo, aceptar esta estruturación, i.e., situar a vogal temática dentro dun tema verbal, implica, cando menos para min, negarlle un valor fundamentalmente flexivo e concederlle un valor primariamente léxico.

Este último punto é o que pretendo demostrar ó longo deste apartado. Para isto, preciso primeiro mostrar que a vogal temática non posúe unha natureza fundamentalmente flexiva e, polo tanto, non debe asociarse primariamente á morfoloxía flexiva (3.1); en segundo lugar, tentarei demostrar a súa natureza primariamente léxica (3.2); e, en terceiro lugar, mostrarei que os usos da vogal temática na morfoloxía flexiva son do mesmo tipo cós usos da vogal radical (3.3).

3.1. A natureza non fundamentalmente flexiva da vogal temática

Segundo afirméi arriba (§ 2.1), considerarei fenómenos flexivos aqueles que teñan repercusións na sintaxe, sexa porque deben ser vistos por regras sintácticas (coma no caso das propiedades inherentes), sexa porque son establecidos pola sintaxe (coma a asignación das propiedades morfosintácticas de concordancia, de configuración ou de frase). Neste sentido, é notorio que o feito de un verbo pertencer á primeira, á segunda ou á terceira conxugación non afecta de ningún xeito á estrutura sintáctica en que se insire, nin ningún elemento da estrutura sintáctica determina nunca a conxugación a que debe pertencer un verbo.

Deste xeito podemos comprobar que non existen posicións sintácticas das que un verbo se poida ver excluído por pertencer a unha conxugación determinada. Tanto os verbos da primeira, coma os da segunda ou os da terceira poden aparecer como V nunha FV que rexa oracións completivas de subxuntivo introducidas por *que*: *mando que veñas, quero que veñas, pido que veñas*; do mesmo xeito, estar adscrito a unha conxugación concreta non lle impide a un verbo funcionar dentro desas oracións subordinadas: *mando que cantes, mando que comas, mando que partas*; calquera verbo en Infinitivo, con independencia da súa conxugación, pode ser introducido por unha preposición: *para cantar isto, para comer iso, para partir iso*. Feitos coma os sinalados mostran que a conxugación non depende da estrutura sintáctica en que se insira o verbo.

Máis aínda, a vogal temática non é un trazo que xere regras de concordancia nin que se vincule con propiedades configuracionais ou de frase: ningún complemento dun verbo, nin ningún suxeito, nin ningún adxunto, nin ningún complementizador (conxuncións coma *que*) ten que concordar de ningún modo coa vogal temática de ningún verbo. Contrástese esta situación coa da concordancia en número: un verbo flexionado para a P1 Singular non pode aparecer como V dunha FV que ten un suxeito P2 Plural: **vós levántome ás seis*. Este tipo de feitos mostra que a conxugación non ten repercusións sobre os outros elementos que se relacionan sintacticamente co verbo.

Máis aínda, non existe ningunha categoría morfosintáctica (modo, tempo, aspecto, voz...) vencellada a esta ou estoutra vogal temática. E, por outra parte, vemos que un verbo non deixa de estar flexionado para determinadas propiedades morfosintácticas por pertencer a unha conxugación calquera: tanto na C-I, coma na C-II e na C-III existen as P1 Plural do Subxuntivo Presente (*cantemos, comamos, partamos*) ou a P3 Singular do Indicativo Presente (*cantou, comeu, partiu*). Con todo, un verbo flexionado para o Infinitivo non pode estar flexionado para a categoría modo, nin para a categoría tempo e, en determinados contextos, para as categorías persoa e número.

Todos estes feitos proban, ó meu modo de ver, que a vogal temática non é un elemento relacionado coa flexión verbal, o que, por outra parte, xustifica a súa inclusión no tema do verbo e a súa exclusión do bloque conformado polos sufixos expoñentes de propiedades morfosintácticas.

3.2. A natureza fundamentalmente léxica da vogal temática

Para poder defender a natureza fundamentalmente léxica da vogal temática debo primeiro recordarlle ó lector que, de acordo coa nosa tradición gramatical, incluíu a vogal temática dentro do tema, o que ó meu xuízo lle nega xa de entrada un carácter primariamente flexivo. Que a vogal temática está dentro do tema queda

fóra de toda dúbida desde o momento en que, como acabo de mostrar, comprobamos que non é un afixo que se axunte á raíz como consecuencia dunha operación flexiva.

O feito de que todo tema verbal galego deba conter unha vogal temática lévame a postular a existencia do principio seguinte na gramática da lingua galega:

- (12) Os temas das entradas léxicas de natureza verbal en galego deben rematar nunha das seguintes tres vogais: /a/, /e/, /i⁵.

Nótese que estou a falar de temas no sentido de formas básicas de palabra da entrada léxica, i.e., da compoñente fonolóxica da entrada léxica. De feito, o principio de (12) é unha *condición de boa formación fonolóxica*⁶ que opera no nivel morfofonémico. Nótese tamén que neste principio non importa o grao de complexidade do tema verbal: a condición vale para temas que só conteñen unha raíz coma *cantar*, *comer* ou *partir*, ou temas complexos coma *enriquecer*.

Defenderei aquí que tanto a vogal temática coma a restrición de (12) son elementos de natureza fundamentalmente léxica, pois atanguen á forma fonolóxica que vai conter unha determinada entrada léxica (recordemos de novo que a forma fonolóxica é un dos elementos compoñentes da entrada léxica). Para acometer esta defensa sen caer na circularidade, debo tamén recordarlle ó lector a división entre flexión e derivación que deseñei arriba. Segundo esta división, a condición de (12) só pode ser considerada de orde léxica, pois como dicía Anderson (vid. § 2.2), “as descrições estruturais das regras derivacionais [*ergo*, das regras de natureza léxica, F.D.] refírense ó contido das (clases de) ítems léxicos”, mentres que as regras flexivas se refíren a propiedades da estrutura morfosintáctica. Unha condición como a de (12) mostra que elementos do repertorio fonolóxico deben estar presentes na estrutura fonolóxica dun tipo de temas, i.e., nun dos compoñentes dunha clase de entradas léxicas. O tema dun verbo debe conter tal elemento con independencia do lugar que este verbo ocupe na sintaxe; e este elemento do tema do verbo nunca inflúe nas outras palabras que gravitan arredor do verbo. Dado que

⁵ De feito, o galego só coñece unha violación desta condición: o verbo *pór* (son conflitivos tamén a este respecto *vir e ter*, véxase Dubert 1999a). Con todo, é preciso facer dúas consideracións: a) *pór* presenta un elemento [ɛ] xeralmente considerado vogal temática (véxase Álvarez e Xove 2002: 255-256) nas formas do Indicativo Pretérito (*puxeches*), do Antepretérito (*puxeras*), do Subxuntivo Pretérito (*puxeses*) e do Futuro de Subxuntivo (*puxer*); b) este verbo mostrou unha tendencia xeneralizada a ir creando formas temáticas noutras partes da súa conxugación, dando orixe ó verbo *poñer*, que segue sendo irregular e que hoxe suplantou a *pór* na maior parte do territorio.

⁶ Trask define as condicións de boa formación nos seguintes termos: “Calquera das varias condicións que deben ser satisfeitas por unha representación fonolóxica, usualmente formulada nos confíns dunha particular teoría da fonoloxía, para que a representación sexa admisible dentro da fonoloxía da lingua en cuestión. Moitas teorías da fonoloxía impoñen varias destas condicións como parte do seu contido” (1996, s. v. *well-formedness conditions*).

a vogal temática cumpre estas dúas condicións, creo que a súa presenza só pode ser explicada como resultado dunha restrición léxica; e dentro das restricións léxicas, unha restrición de carácter fonolóxico, pois afecta á compoñente fónica da entrada léxica: á secuencia de fonemas, /kome/, que se coindexan cun significado conceptual, COMER, e cunha etiqueta sintáctica V_1 , como mostrei en (7). Esta idea xa a expresou, como indiquei máis arriba, Porto Dapena (1972).

Tomemos unha raíz como /kaNt/, que aparece na lingua tanto na entrada léxica *cantar* (V) coma na entrada léxica *canto* (Subs). Se queremos empregar a entrada léxica *cantar*, debemos conformar o seu tema /kaNt/ sumándolle unha vogal temática, neste caso /a/, á raíz /kaNt/. A engádegas da vogal temática é unha operación fonolóxica que permite construír o tema dunha entrada léxica verbal. O mesmo sucede con temas máis complexos, coma /eNrikeθe/, tema da entrada léxica *enriquecer*. O seu tema /eNrikeθe/ está composto por unha raíz /rik/ á que se lle engade un prefixo /eN/ e un sufixo /eθ/; esta terminación, que se emprega como afixo en verbos relacionados con adxectivos (*ennegrecer*, *embrutecer*, *engrandecer*, *palidecer*) e aparece como terminación de verbos que non se relacionan con ningunha base illada (*parecer*, *perecer*, *coñecer*), selecciona unha vogal temática /e/.

A escolla da vogal é moitas veces un feito idiosincrásico, produto, sobre todo, de factores históricos. A entrada léxica *bater* toma unha vogal temática /e/, mentres que *matar* toma unha vogal temática /a/ e *partir* toma unha vogal temática /i/. Non hai nada en *bater* que nos obligue a empregar a vogal /e/. A condición establecida en (12) simplemente estipula que todo tema verbal debe ter unha vogal temática, e que esta vogal só pode ser /a/, /e/ ou /i/. Con todo, este comportamento idiosincrásico só existe para a vogal radical /a/, pois noutros casos parece existir algún tipo de restricións fonotáticas que sinalan certas preferencias na escolla:

A CI é a máis productiva e numerosa de tódalas conxugacións, pois máis das tres cuartas partes dos verbos intégranse nesta clase conxugacional. Esta amplitude ocorre tamén na VR [vogal radical, F.D.], pois é a única conxugación que pode ter de VR calquera vogal, e incluso ditongos: *choutar*, *amainar*, *estreitar*... Nos verbos da CII e da CIII, en cambio, a VR actúa tamén como marca de clase: as VR pechadas [i] e [u] só existen na CIII, non na CII (Álvarez e Xove 2002: 247).

Polo tanto, as vogais radicais altas /i/ e /u/ son compatibles coa vogal temática /a/ (*pisar*, *puxar*) e coa vogal temática /i/ (*vivir*, *durmir*), pero incompatibles coa vogal temática /e/. De feito, ó longo da historia do galego deuse un proceso fonolóxico diacrónico que tendeu a harmonizar a vogal temática /e/ de verbos coma *viver*, *dicir* coa radical, dando orixe a *vivir* e *dicir*. Alén disto, e como sinala Santamarina (1974: 14) a resultas da neutralización de /e/ e /i/ postónicos (véxase aquí en § 3.3.2), as vogais radicais /e/ e /o/ tenden a preferir a vogal temática /e/. Os verbos

da C-III con vogal radical /e/ mostran tendencia a manter comportamentos atípicos na súa vogal radical (*pido, pides, pedimos...*) e practicamente non existen verbos da C-III con vogal radical /o/. Con todo, non é este lugar para entrar nesta apaixonante discusión, pois a finalidade deste traballo é mostrar o carácter léxico da vogal temática dos verbos.

3.3. O uso flexivo da vogal temática

A forma final dunha forma de palabra depende de varios factores, que dan orixe a outros tantos padróns de alternancia (inspírome en Mohanan 1995):

- α) Os padróns de alternancia alofónica, que se refiren ás variacións na forma da palabra motivadas pola alternancia de alófonos. Unha característica básica e obrigatoria deste tipo de alternancia é que está condicionada por factores estritamente fónicos. Un exemplo típico é o que lle acontece á nasal posnuclear en galego: unha nasal na coda dunha sílaba presentará o punto de articulación da primeira consoante do ataque da sílaba seguinte.
- β) Os padróns de alternancia fonémica, que se producen cando vemos alternar cadeas de fonemas por condicionamentos de orde estritamente fónica. Un exemplo claro constitúeo o seseo posnuclear dos dialectos galegos noroccidentais, onde encontramos /s/ en posición posnuclear, *pas*, e /θ/ en posición prenuclear, *paces* (véxase Dubert 1999a e 1999b).
- γ) Os padróns de alternancia morfofonémica, en que alternan as secuencias de fonemas en contextos nos que se mesturan condicións de tipo fónico con outras sintácticas ou léxicas. Compárese o comportamento das vogais radicais de *pedir* co comportamento das vogais radicais de *seguir* (véxase Dubert 1999a ou Álvarez e Xove 2002).
- δ) Os padróns supletivos, nos que, fronte ós tres anteriores, non é posible encontrar ningún tipo de relación formal sistemática e predicible entre unhas formas de palabra e outras: é o caso da conxugación de verbos coma *ser* (*eu son, ti es, eu fun, eu era*) ou *ir* (*vou, imos*). Na análise que aquí propoño, as formas supletivas están dalgún xeito almacenadas no lexicón en entradas léxicas que xa conteñen, cando menos, algunhas das propiedades morfosintácticas que determinan as formas⁷.

Serán os padróns de (β) e de (γ) os que me interesa explorar aquí porque permiten explicar o comportamento da vogal temática durante a flexión dos verbos, pois

⁷ A súa selección depende da condición *nos outros sitios* ('elsewhere condition') que xa vimos funcionar na selección dos alomorfos do artigo definido e do pronome acusativo de terceira persoa (véxase Dubert 2001).

nese apartado debo dar conta dos casos de desaparición da vogal temática na flexión dos verbos regulares e da aparición de formas de vogal temática diferentes das que aparecen no tema da entrada léxica: todos estes procesos serven para que un elemento de natureza léxica adquira funcións flexivas secundarias.

3.3.1. A desaparición da vogal temática

Como é sobradamente sabido, a P1 Singular do Indicativo Presente (*cant-o*, *com-o*, *sub-o*) e de todo o Subxuntivo Presente (*cant-e*, *com-a*, *sub-a*) presenta formas de palabra atemáticas nos verbos regulares das tres conxugacións. Este feito podería constituír un contraexemplo á condición de boa formación que presentei en (12) e que estipulaba que o tema dos verbos debe ter unha vogal temática. Podería, ademais, interpretarse como un indicio da natureza fundamentalmente flexiva da vogal temática, pois da presenza ou ausencia desta resulta a aparición de formas de palabra para diferentes palabras morfosintácticas. A forma de palabra da P1 Singular do Indicativo Presente diferénciase da forma de palabras da P3 Singular do mesmo tempo e modo en que na primeira aparece o sufixo número-persoal /o/ e na segunda a vogal temática /a/: /kaNto/ vs. /kaNta/; e as P2 Singular, P3 Singular, P1 Plural, P2 Plural e P3 Plural do Indicativo Presente diferénciase das formas de palabra correspondentes do Subxuntivo Presente en virtude da presenza nas primeiras de vogal temática /a/ e de sufixo modo-tempo /e/ nas segundas: /kaNta/ vs. /kaNte/. A vogal temática posuiría así un valor primordialmente flexivo igual ó que teñen os sufixos de número-persoa /o/ e os sufixos de modo-tempo /e/~/a/.

Con todo, estes problemas quizais só sexan aparentes. Existe unha longa tradición que proclama a elisión, dentro dun proceso sincrónico actual, da vogal temática nestes contextos (véxase por exemplo Mateus e Andrade 2000); entre os nosos lingüistas, tamén Santamarina (1974: 15), que emprega un modelo descritivo de *ítem e proceso* (véxase Pena 1990) máis ou menos similar ó que uso aquí, insinuou que “en la P¹ de las tres conjugaciones la vocal temática es Ø. Desde el punto de vista sincrónico resulta difícil explicar la razón de su ausencia a no ser que la interpretemos como una asimilación al sufijo modo-temporal [sic, F.D.] /-o/”. Se así fose, a vogal temática estaría presente no nivel morfofonémico en que opera (12) e desaparecería nun nivel máis baixo como resultado dun principio fonolóxico do tipo (γ)⁸.

⁸ As formas do P1 Singular do Indicativo Presente e de todo o Subxuntivo Presente constitúen un lugar de procesos flexivos comúns que as individualizan: compárese *vallo ~ valla* con *vales*, *fago ~ faga* con *fas*, *sirvo ~ sirva* con *serve*, *caibo ~ caibas* con *cabes*, *roio ~ roia* con *roes*, e un longo etcétera (Dubert 1999a). Ademais, a escolla de /e/ como sufixo de modo-tempo na C-I e de /a/ como sufixo de modo-tempo na C-II e C-III podería estar condicionada fonoloxicamente. A vogal /a/ é *non-tensa* mentres que as vogais /e/ e /i/ son *tensas*; a vogal temática /a/ *non-tensa* escolle como alomorfo de modo-tempo unha vogal *tensa* /e/; as vogais temáticas *tensas* /e/ ~ /i/ escollen como alomorfo de modo-tempo unha vogal *non-tensa* /a/. Se este condicionamento fonolóxico fose real, mostraría que as vogais temáticas deben estar presentes nalgún momento e influen na escolla dos alomorfos do sufixo de modo-tempo do Subxuntivo Presente.

Se isto fose así, a vogal temática adquiriría valor flexivo de xeito secundario como consecuencia da súa desaparición na P1 Singular do Indicativo Presente e en todo o Subxuntivo Presente. Con todo, a súa presenza inicial no tema verbal nun primeiro momento deberíase á restrición léxica de (12).

3.3.2. Variacións na forma da vogal temática

Por outra parte, é un feito que cando aparece a vogal temática nunha forma de palabra, esta non é sempre a mesma que figura no tema da entrada léxica:

- (13) a) Na C-I a vogal temática é sempre /a/ salvo na P1 Singular do Indicativo Pretérito, en que aparece /e/: *cantei*; e na P3 Singular do Indicativo Pretérito, en que aparece /o/: *cantou*.
- b) Na C-II a vogal temática é sempre /e/ salvo nas formas de palabra do Copretérito, das P1 Singular e P2 Singular do Indicativo Pretérito e do Participio, a vogal temática é /i/: *collía*, *collín*, *colliches*.
- c) Na C-III a vogal temática é sempre /i/ salvo nas P2 Singular, P3 Singular e P3 Plural do Indicativo Presente e na P2 Singular do Imperativo, en que figura /e/.

Os cambios de (13a) explícanse por medio de principios morfofonémicos (principios do grupo γ): na forma de palabras da P1 Singular, a vogal temática /a/ do nivel morfofonémico correspóndese cunha vogal no nivel fonémico que comparte os trazos [-posterior, -baixo] do sufixo de número-persoa /i/, mentres que no caso da forma de palabras da P3 Singular a vogal temática /a/ do nivel morfofonémico correspóndese cunha vogal no nivel fonémico que comparte os trazos [+posterior, -baixo] do sufixo de número-persoa /u/. Estas asimilacións, moi comúns en galego noutros contextos, aquí son obrigatorias. O resultado desta asimilación volve a ter utilidade flexiva, pois reforza as marcas de P1 Singular e P3 Singular. Pero este valor flexivo adquirese grazas a que a vogal debe estar presente no nivel morfofonémico en virtude da condición léxica de (12).

Pola súa parte, os cambios de (13b) e (13c) poden ser tamén explicados a través de principios morfofonémicos:

- (14) a vogal temática /e/ no nivel morfofonémico correspóndese con /i/ no nivel fonémico cando /e/ vai seguida do expoñente /a/ de [Copretérito]⁹.

⁹ Nótese que /a/ é un alomorfo en distribución complementaria e que a condición que goberna a escolla de /ba/ ou /a/ como expoñentes de Copretérito pode ser formulada en termos puramente fonolóxicos: escóllese /ba/ cando o tema remata en vogal /a/ e escóllese /a/ cando o tema remata en vogal /e/ ou /i/.

Con todo, fai falla sinalar que non existe ningún condicionamento na fonoloxía automática actual da lingua galega que conspire contra o hiato /ea/ no nivel fonémico: *lea, leamos, cadea...*, aínda que este condicionamento xa existiu antes na historia do galego (lat. vul. *DĪA > *[de]a > gal. [d]ia ‘día’). Polo tanto, as formas tipo *collía* contrastan coas formas tipo *andei* e *andou* de (13a) en que estas últimas posúen unha base fonolóxica máis natural. Polo que se refire ás formas tipo *collín* e *colliches*, poderían ser explicadas do mesmo xeito:

- (15) unha vogal [-alta, -baixa] ó final do tema verbal no nivel morfofonémico correspóndese cunha vogal [+alta] no nivel fonémico cando vai seguida dos expoñentes /N/ de [P1, Singular, Indicativo, Pretérito] ou /t̪es/ de [P2, Singular, Indicativo, Pretérito].

O principio (15) vale tanto para *comín* e *comiches*, fronte a *comeu, comemos, comestes, etc.*, dun lado e *cantaches* e *partiches* doutro; como para *fun* e *fuches*, fronte *fomos, fostes, foras*. Como xa sinaléi antes (Dubert 1999a), as formas da P1 Singular e P2 Singular do Indicativo Pretérito de *ser* e *ir* compórtanse a este respecto coma as formas dos verbos regulares da C-II (incluso parece común que os falantes que empregan *fumos* empreguen tamén *comimos*).

Con todo, estas propostas de (14) e (15) levantan máis problemas dos previstos. Así por exemplo, en primeiro lugar, (15) fala de /N/ como expoñente de [P1, Singular, Indicativo, Pretérito] e de /t̪es/ como expoñente [P2, Singular, Indicativo, Pretérito], cando /N/ e /t̪es/ son comunmente vistos como sufixos de número-persoa. En segundo lugar, e máis importante, os conxuntos de propiedades [Copretérito], [P1, Singular, Indicativo, Pretérito] e [P2, Singular, Indicativo, Pretérito] non só se expresan na forma de palabra a través de /a/, /N/ e /t̪es/, respectivamente, senón que *tamén* se expresan por medio da elevación da vogal: *comín* fronte a *comemos* e *comiches* fronte a *comestes* (véxase 2.4); *comíamos* fronte a *comemos*. Polo tanto, cómpre cuestionarse se a elevación das vogais en (14) e (15) se debe á contigüidade dos devanditos expoñentes ou se pode ser explicada directamente pola presenza das propiedades morfosintácticas. Desde un punto de vista histórico, a aparición da nasal no caso da P1 Singular como expoñente de número-persoa é posterior á elevación da vogal (a vogal tónica xa era /i/ antes de que se inserise /N/, véxase Santamarina 1974: 55-56 e 64), mentres que a elevación na P2 Singular débese á metafonía que exerceu un antigo /i:/ latino que aparecía na concatenación de afixos -īs-tī que deu orixe ó galego -ches (véxase Santamarina 1974: 55). Moitos procesos morfofonémicos que se adoitan postular como sincrónicos son case sempre paralelos ós cambios históricos que dan orixe ás formas de palabra que se quere analizar. Por esta razón podemos falar de que o expoñente /a/ de Copretérito eleva a vogal temática /e/ a /i/ (véxase Santamarina 1974: 53). Con todo, o que nos interesa é a explicación sincrónica das alternancias, porque os falantes non teñen unha gramática histórica na cabeza e o que se

pretende na descrición sincrónica é mostrar o coñecemento actual que os falantes teñen da lingua; polo tanto, a explicación sincrónica non ten por que seguir a diacrónica. Polo de agora, non me pododo decidir nesta cuestión. En principio direi, provisoriamente, que a vogal temática /e/ no nivel morfofonémico correspóndese coa vogal /i/ no nivel fonémico cando V está determinado polos conxuntos de propiedades morfosintácticas a que se refiren (14) e (15)¹⁰.

Outra vez, son os cambios que padecen as vogais temáticas os que teñen utilidade na morfoloxía flexiva, pero non as vogais en si, que seguen sendo requiridas pola condición de (12). Este uso flexivo vese máis claramente no caso de *fun ~ fuches*, formas atemáticas en que é a alternancia da vogal radical o que ten valor flexivo, sen que a vogal radical perda por iso a súa natureza léxica como elemento constituínte do tema verbal.

Finalmente, no caso de (13c) observamos simplemente como actúan regras de alternancia fonémica do tipo (β). Debido a outra condición de boa formación fonolóxica, o segmento /i/ non pode ocupar un núcleo silábico postónico, polo que fai falla unha regra fonolóxica de reaxuste que adapte o resultado de (12) ás restricións fonolóxicas de superficie propias da lingua galega:

- (16) /i/ no nivel morfofonémico correspóndese con /e/ no nivel fonémico cando ocupa un núcleo silábico en posición postónica.

É claro que o principio de (16) é enteiramente fonolóxico. Como sinala outra vez Antón Santamarina:

El comportamiento de la vocal temática está comprometido por el acento. En aquellas personas en que es tónica (P⁴, P⁵) la vocal aparece inconfundible. En los casos en que es átona (P², P³, P⁶) sólo se mantiene en la CI puesto que /a/ aun en posición átona no se neutraliza con ninguna otra vocal. En las CII y CIII las vocales temáticas /e/ e /i/ se neutralizan ambas en /e/ (1974: 14).

Polo tanto, e outra vez, o valor flexivo das alternancias /i/ ~ /e/ nos verbos regulares da C-III (e.g. *partes* vs. *partimos*) non se debe á existencia da vogal temática, que é un requirimento de (12), senón que neste caso é o resultado de procesos fonolóxicos regulares. No fondo, a alternancia *partes* ~ *partimos* é semellante á que observamos en *pas* ~ *paces* nos dialectos de seseo posnuclear: nestas parellas, o

¹⁰ Realmente, estamos presenciando cómo decrece a naturalidade fonolóxica dos procesos morfofonolóxicos. Ordenados de máis a menos naturais, temos que o cambios /ai/ > /ei/ e /au/ > /ou/ son os máis motivados na fonoloxía sincrónica do galego; pola súa parte, o cambio /ea/ > /ia/ está menos motivado; finalmente, a elevación /e/ > /i/ de *comín* ~ *comiches* e de /o/ a /u/ en *fun* ~ *fuches* carece de toda motivación fonolóxica sincrónica.

segmento fricativo dental alterna cun segmento fricativo alveolar en virtude de condicións estritamente fonolóxicas; aínda que esta alternancia permite distinguir a forma de palabra do singular *pas* da forma de palabra do plural *pac-es*, non ten unha motivación flexiva, senón que resulta dun proceso condicionado pola fonoloxía actual desas variedades lingüísticas, e esta alternancia pode ser empregada/interpretada para facilitar a diferenza entre Singular e Plural.

4. Conclusións

Neste traballo tentei demostrar o carácter primordialmente léxico da vogal temática dos verbos. Para iso, tiven primeiro que mostrar que a súa inserción non se debe a un proceso flexivo, pois non é un elemento requirido polas estruturas sintácticas en que se insire o verbo nin a súa inserción repercute en ningunha das unidades que se enxertan nas expansións sintácticas do verbo (sexan frases verbais, sexan oracións); de feito, a vogal temática non se relaciona con ningunha propiedade morfosintáctica.

En segundo lugar, tiven que mostrar que os usos na flexión da vogal temática son secundarios e indistintos, por exemplo, dos usos flexivos da raíz dos verbos, elemento xeralmente relacionado máis coa compoñente léxica ca coa sintáctica.

En terceiro lugar, tiven que mostrar que tódolos comportamentos anómalos das vogais temáticas (tanto a súa ausencia coma as súas alternancias formais dentro de cada paradigma) poden ser explicados a partir de diferentes tipos de procesos fonolóxicos, nomeadamente, procesos morfofonémicos ou procesos fonémicos. Incluso a propia existencia da vogal temática resulta dunha condición de boa formación fonolóxica dos temas verbais, que requiren ser rematados en /a/, /e/, /i/, de xeito que a vogal temática nin é o morfo de ningún morfema nin o significante de ningún signo, senón un mero formante do significante do signo en que aparece.

En definitiva, todas estas razóns mostran ó meu xuízo a natureza primariamente léxica da vogal temática dos verbos galegos.

Bibliografía

- Álvarez, R. / Xove, X. (2002): *Gramática da lingua galega* (Vigo: Galaxia).
- Álvarez, R. / Monteagudo, H. / Regueira, X. L. (1986): *Gramática galega* (Vigo: Galaxia).
- Anderson, S. R. (1988): "Morphological theory", en Newmeyer, F. J. (ed.): *Linguistics: the Cambridge survey*. Vol. 1. *Linguistic theory: foundations*: 146-191 (Cambridge: CUP).

- Anderson, S. R. (1992): *A-morphous morphology* (Cambridge: CUP).
- Beard, R. (1998): “Derivation”, en Spencer, A / Zwicky, A. (eds.): *The handbook of morphological theory*: 44-65 (Oxford: Blackwell).
- Dubert García, F. (1999a): *Aspectos do galego de Santiago de Compostela* (Santiago de Compostela: Universidade).
- Dubert García, F. (1999b): “Máis sobre o rotacismo de /S/ en galego”, en Álvarez, R. / Vilavedra, D. (coords.): *Cinguidos por unha arela común. Homenaxe ó profesor Xesús Alonso Montero*, Vol. 1: 367-387 (Santiago de Compostela: Universidade) (Edición dixital en: <http://web.usc.es/~fgdubert/rotacism.pdf>).
- Dubert García, F. (2001): “A alomorfía do artigo definido galego á luz da fonoloxía prosódica”, en Regueira, X. L. / Veiga, A.: *Da gramática ó dicionario. Estudos de lingüística galega*: 91-105 (Santiago de Compostela: Universidade). (Edición dixital en: <http://web.usc.es/~fgdubert/prosarti.pdf>).
- Fernández Rei, F. (1998): “Vocal temática e conxugación verbal. A expresión da clase mórfica en español, catalán e galego moderno”, en Kremer, D. (ed.): *Homenaxe a Ramón Lorenzo*, Vol. 2: 591-613 (Vigo: Galaxia).
- Freixeiro Mato, X. R. (2000): *Gramática da lingua galega. III. Morfosintaxe* (Vigo: A Nosa Terra).
- García-Miguel, J. M. / C. Cabeza (2000): “Sintaxe”, en Ramallo, F. / Rei-Doval, G. / Rodríguez Yáñez, X. P.: *Manual de ciencias da linguaxe*: 543-584 (Vigo: Xerais).
- González Seoane, E. X. (1999): “A derivación verbal isocategorial en galego”, en Álvarez, R. / Vilavedra, D. (eds.): *Cinguidos por unha arela común. Homenaxe a Xesús Alonso Montero*. Vol. 1: 473-500 (Santiago de Compostela: Universidade).
- Jackendoff, R. (2002): *Foundations of Language* (Oxford: OUP).
- Martinet, A. (1991): *Éléments de linguistique générale* (Paris: Colin).
- Mateus, M. H. Mira / Andrade, E. d’ (2000): *The phonology of portuguese* (Oxford: OUP).
- Matthews, P. H. (1991): *Morphology* (Cambridge: CUP).
- Mohanan, K. P. (1995): “The organization of the grammar”, en Goldsmith, J.: *The handbook of phonological theory*: 24-69 (Oxford: Blackwell).
- Pena, J. (1990): “Sobre los modelos de descripción en morfología”, *Verba*, 17: 5-75.
- Pena, J. (1991): “La palabra: estructura y procesos morfológicos”, *Verba*, 18: 69-128.

- Porto Dapena, X. Á. (1972): “Encol da estrutura do verbo galego”, *Grial*, 35: 13-28.
- Santamarina, A. (1974): *El verbo gallego* (Santiago de Compostela: Universidade).
- Stump, G. T. (1998): “Inflection”, en Spencer, A. / Zwicky, A. (eds.): *The handbook of morphological theory*: 13-43 (Oxford: Blackwell).
- Trask, R. L. (1993): *A dictionary of grammatical terms in linguistics* (London: Routledge).
- Trask, R. L. (1996): *A dictionary of phonetics and phonology* (London/New York: Routledge).
- Trask, R. L. (2000): *The dictionary of historical and comparative linguistics* (Edinburgh: Edinburgh University Press).

Armando Cotarelo Valledor, editor e estudioso do romanceiro tradicional galego

José Luís Forneiro
Universidade de Santiago de Compostela

Resumo:

Armando Cotarelo Valledor tentou converter-se na maior figura intelectual do galeguismo da sua época, com este fim apropriou-se dalguns textos romancísticos recolhidos pelo advogado e folclorista Alfonso Hervella Courel. Na manipulação e nas glosas destes romances transparece a ideologia ultraconservadora deste filólogo de inícios do século XX.

Palabras chave:

Romanceiro. História da literatura galega. História do nacionalismo galego.

Abstract:

Armando Cotarelo Valledor, a philologist of the beginning of XX c., tried to become the most important intellectual figure of the galeguismo of his time. With this intention he appropriated some of the ballads, which were collected by the lawyer and folklorist Alfonso Hervella Courel. The ultraconservative ideology of this philologist is evident in the manipulation and the notes of these ballads.

Key words:

Collection of ballads. History of the Galician literature. History of the Galician nationalism.

Armando Cotarelo Valledor (Veiga do Eu, 1879-Madrid, 1950) provavelmente seja, na actualidade, uma das figuras menos interessantes da cultura galega moderna. A sua erudição *rancia*, a artificialidade da sua obra literária, o facto de ter sido redigida em língua castelhana a maior parte da sua produção intelectual, a sua adesão ao golpe militar de 1936 e, até, o seu físico pouco atraente, fazem com que este vulto do galeguismo das primeiras décadas do século XX mal interesse hoje ao mundo cultural galego. No entanto, desde a sua chegada a Santiago de Compostela em Outubro de 1904, foi uma das mais singulares figuras públicas na Galiza do seu tempo, pois actuou como um decidido renovador da estagnada universidade compostelana de princípios do século passado (Barreiro 1984: 5-8), e este espírito modernizador não se limitou às aulas e à política universitária, uma vez que Cotarelo se destacou na vida social de Compostela pelo seu convívio com

os estudantes (Filgueira 1984: 12) e até porque se deslocava pela cidade apostólica de bicicleta (Carvalho Calero 1984: 11). Mas Armando Cotarelo foi conhecido, principalmente, na Galiza das primeiras décadas do século XX, por ter participado nalgumas empresas políticas¹ e, sobretudo, como animador da vida cultural, já que, entre outros cargos, foi secretário da Exposição Regional de 1909, presidente do Ateneu León XXIII e do Seminário de Estudos Galegos² e director das revistas *El Centenario* e *Utreya*.

Cotarelo Valledor, na sua estada em terras galegas, que decorreria durante mais de três décadas (1904-1939), ocupou-se, nomeadamente, de contribuir para a produção teatral em língua galega e também de estudar diversos aspectos da história e da literatura da Galiza. A literatura popular não foi alheia aos seus interesses, destarte, não só editou ou dedicou páginas de estudo às composições literárias com que o povo galego acompanhava os seus momentos de labor e de lazer (o conto, a lírica e o romance tradicional), como também incorporou este tipo de literatura às suas obras de criação.

Em 1919, Cotarelo publica na *Utreya*, “Revista Quincenal de Cultura Galaica”, um artigo intitulado “Romancero popular gallego”, onde pretendia

recoger en estas páginas cuantos romances de Galicia lleguen a nuestras manos, pues de este modo pensamos servir de alguna utilidad a quien haya de emprender el estudio y publicación del *Romancero popular gallego* que falta en nuestra región. [...]

Existe, sin embargo, caudal de este curioso género literario, viviente en boca del pueblo gallego, y de él poseemos algunas muestras que deseamos acrecer con las comunicaciones de amigos, discípulos y amantes de Galicia. Todas serán incluidas en la Revista, expresando la procedencia y nombre del colaborador (Cotarelo 1919: 33).

Uma vez manifestada a sua intenção de dar a lume quantos textos do acervo romancístico pudessem chegar às suas maos, Cotarelo julgava que

Aunque todavía consideramos prematuro, a lo menos para nosotros, el análisis de los romances galaicos, no debemos omitir ligeras consideraciones generales sobre ellos, para ilustración y recordanza de cuantos gusten favorecer esta empresa con sus aportamientos y sus luces (Cotarelo 1919: 33).

¹ Foi eleito deputado pelo partido liberal em 1919 (Barreiro 1984: 8) e organizou uma frente eleitoral nos inícios da República de 1931 (Fernán Vello e Pillado Mayor, 1986: 46).

² Membros do Seminário de Estudos Galegos como Ricardo Carvalho Calero, José Filgueira Valverde e Ramón Martínez López reconheceram o magistério intelectual de Cotarelo na instituição.

Como vemos, o nosso autor pretendia aqui ofrecer uma visão geral do romanceliro galego e ao tempo converter-se na máxima autoridade na matéria. Assim, Armando Cotarelo dava conta das colecções e das publicações realizadas até ao momento, e informava, de modo esquemático, sobre diversos aspectos do romance na Galiza como o nome, a classificação (pela linguagem, pela origem, pelo assunto), o mérito, o número, a datação, o assunto, os títulos, a transmissão, o estado e a música. Nas duas páginas em que se ocupava de todas estas questões, Cotarelo só se mostrava em parte devedor das teses de Manuel Murguía, o maior estudioso deste género na Galiza do século XIX. Este afirmara inicialmente que o romanceliro era um género literário pouco conhecido na Galiza, para passar poucos anos depois a reconhecer a sua vitalidade e riqueza, chegando a afirmar que, para além dos romances de procedência castelhana, os de versos de oito sílabas, o saber folclórico galego atesourava canções narrativas mais antigas, baladas de origens celta e germânica, em versos hexassilábicos e heptassilábicos. Do marido de Rosalia de Castro, Cotarelo fez próprias ideias como a da pobreza numérica e poética do romanceliro face a outros géneros tradicionais e a da escassa presença na tradição da Galiza do verso romancístico por excelência, o verso octossilábico. No entanto, no que se refere ao aparecimento do romanceliro na Galiza, don Armando desentendia-se de improváveis origens pré-romanas ou suevas para assumir a sua importação do centro peninsular e assim, afirmava, baseando-se em Menéndez Pelayo, que os romances derivaram dos velhos cantares de gesta da Castela de fins do século XIV ou de princípios do século seguinte, e os mais antigos provavelmente chegaram à Galiza no século XV (Cotarelo 1919: 34).

Na realidade, esta como as outras informações e afirmações contidas neste trabalho (se exceptuarmos as já referidas de Murguía e as relativas à música dos romances) estavam tomadas da introdução que Alfonso Hervella Courel redigiu para o seu *Romances populares gallegos recogidos de la tradición oral*. Este advogado, nascido em Viana do Bolo em 1886³, apresentara em 1909 a sua colecção de romances, obtida na sua terra natal, à *Exposición Regional Gallega*, de que Armando Cotarelo era secretário. Porém, este no seu artigo “Romanceliro popular gallego” não cita em nenhum momento o trabalho do estudioso ourensano como fonte informativa. De facto, na segunda nota de rodapé que aparecia em *Romanceliro popular gallego*, as palavras de Cotarelo pareciam indicar que já não estava em seu poder o romanceliro de Hervella:

³ Expresso aqui o meu agradecimento a Alfonso Mato, do Seminário de Estudos Galegos, e a Álvaro Armesto, juiz de paz de Viana do Bolo, as informações que tão atentamente me forneceram sobre este esquecido vianês.

En la Exposición regional gallega celebrada en Santiago en 1909, tuvimos el gusto de examinar este conjunto de romances, todos recogidos de la tradición oral de la provincia de Orense. El autor residía entonces en Viana del Bollo a donde se le devolvió el manuscrito y es hoy dignísimo notario de Monforte (Cotarelo 1919: 33).

Resulta sospeitoso que Cotarelo indicase neste artigo que todas as versións que lle fosesen enviadas “serán incluídas en la Revista, expresando la procedencia y nombre del colaborador”, e que no caso de Hervella chegase a proporcionar una información, em principio desnecesaria: que o manuscrito que ele conhecera como secretario da Exposición de 1909, lle fora devolvido ao seu autor, todo “un dignísimo notario” da vila de Monforte de Lemos. Parece que Cotarelo, valendo-se do seu papel social e intelectual na Galiza do seu tempo, pretendia ser o “coordinador xeral” na edición do romanceiro galego que Alfonso Hervella demandava no prefacio da súa colectánea:

Yo espero que algún literato pacienzudo y discreto recoja y seleccione aquellos de nuestros mejores romances que están esparcidos por las cuatro provincias gallegas y forme con ellos una más completa colección que esta que hoy ofrezco a mis paisanos. La labor es delicada y merece que anden en ella manos expertas y cuidadosas (Hervella 1909: II-III).

Daí que Don Armando expressasse claramente a súa vontade de proclamar-se como o maior especialista no romanceiro galego no final do seu artigo:

Sin duda que la presente colección debía de sujetarse a algún orden, adoptada una clasificación previa de los romances gallegos; mas para esto sería preciso conocerlos todos o por lo menos en su mayor parte, y aunque poseemos no despreciable caudal, fáltannos seguramente muchísimos, que iremos sumando con el favor de los entusiastas, peritos y amantes de Galicia a quienes, ahora y siempre, suplicamos ayuda y consejo. Los daremos, pues, a la publicidad conforme nos sea posible y lleguen a nuestro conocimiento, ya que aun desordenada, esta colección ofrecerá la ventaja de presentar reunido lo que anda disperso y de hacer asequible lo inédito o poco notorio. Más tarde vendrá el clasificarlos, estudiarlos y disponerlos para la edición crítica y científica de tan curiosas manifestaciones del arte popular. Entretanto acopiemos materiales para ello (Cotarelo 1919: 36).

São, portanto, sospeitas estas *excusationes non petitae* sobre as orixens das súas fontes, que se contradizem com o ocultamento de que a maioría das informacións de “Romancero popular gallego” estaba contida na introdución do romanceiro inédito de Hervella, para além de nunca ter sido antes formulada por ninguém. Estas suposicións toman maior forza se confrontarmos os textos publicados por Cotarelo com as versións da colección de Hervella.

Quando da altura da realização da nossa tese de doutoramento, *El bilingüismo en el romancero tradicional gallego*,⁴ tivemos que analisar o romance de *Tamar* para ver qual a presença da língua galega neste tema, reparámos, após a leitura das duas versões recolhidas por Alfonso Hervella, que o texto deste romance publicado por Armando Cotarelo na revista *Nós* fora fabricado com as duas versões do primeiro devidamente retocadas. Tal descoberta levou-nos a comparar os romances publicados por Cotarelo, em que ele aparecia como colector, com as versões da colectânea de Hervella e, em praticamente todos os casos, verificámos que os textos dados a lume pelo erudito da Veiga do Eu também eram devedores dos romances obtidos por Hervella Courel.

Ignoramos se, realmente, lhe foi restituída a Alfonso Hervella a sua colecção de romances, mas muito provavelmente a cópia de que fala Armando Cotarelo é a que se conserva hoje no Museu de Ponte Vedra. Esta cópia é praticamente igual a que enviou a Ramón Menéndez Pidal, estando ambas datadas de 10/III/1909, ou seja, do ano em que se celebrou, em Santiago de Compostela, a *Exposición Regional Gallega*.⁵ Existe uma terceira cópia, um tanto diferente das anteriores, que foi a que Hervella Courel preparou como trabalho de ingresso no Seminário de Estudos Galegos, em 25 de Abril de 1927. Nesta terceira cópia, Hervella reduzia sensivelmente o estudo introdutório e o *corpus* textual por ele recolhido, bem como retocava os materiais romancísticos que obtivera da tradição oral (que nalguma ocasião atribuía a lugares donde não procediam os textos), não indicava o nome dos informantes e incorporava versões publicadas por outros autores, como António de la Iglesia. Por razões de acesso aos materiais podemos supor que Cotarelo conheceu as duas colecções depositadas nas instituições galegas, porém o facto de a maioria dos romances publicados por Cotarelo ver a luz antes do ingresso de Hervella no Seminário de Estudos Galegos, assim como as características dos textos publicados, é prova determinante de que Armando Cotarelo manuseou a cópia que se encontra hoje no Museu de Ponte Vedra, como veremos a seguir.

Para o confronto entre as versões obtidas por Hervella e as editadas por Cotarelo oferecemos como texto base os romances pertencentes à colecção do Museu de Ponte Vedra, indicando em notas de roda-pé as variantes que correspondem às cópias do Archivo Menéndez Pidal (AMP) e do Seminário de Estudos Galegos (SEG). Para mostrar com que versões da colecção de Hervella elaborou Cotarelo os seus textos, identificaremos cada uma delas com uma letra que antecederá o verso, hemistíquio ou partes do hemistíquio editados por Cotarelo. As variantes introduzidas por este são

⁴ Para tal fim, tivemos que estudar as mais de 6.000 versões dos mais de cento e sessenta temas que conformam o *corpus*, na sua maioria inédito, do romanceliro da Galiza.

⁵ As diferenças entre as duas colecções são mínimas e estão originadas pelo seu carácter manuscrito que inevitavelmente acarreta algumas divergências de cópia. A colectânea depositada na instituição pontevedresa faz parte do espólio de Casto Sampedro (IV Manuscritos: 5) (Filgueira 1988: 307).

marcadas em itálico e respeitamos a ortografía e a puntuación orixinais de ambos os editores. Após evidenciar como o erudito da Veiga do Eu fabricou os seus romances, trataremos dos comentários que os acompañavam, dado que são muito significativos de qual era a súa visión do romance tradicional e da cultura galega.

No artigo “Romancero popular gallego” Cotarelo completava o seu estudo con una versión do tema intitulado polo Seminario Menéndez Pidal no *Índice General del Romancero (IGR) Una fatal ocasión*, a que Hervella tinha dado o título de *A romeira*, título conservado por Cotarelo.

A romeira I (a)

Rubiales (Viana del Bollo). Petra Fernández.

- Por aqueles campos verdes linda romeira camiña.
 2 Paseando s’iba, paseando ô estilo d’Andalucía;
 peinando s’iba, peinando c’un peine de prata fina.
 4 Saya leva sobre saya; basquiña sobre basquiña;
 zapato leva picado que de lonxe relucía;
 6 as medias leva de lana que parecen prata fina.
 Mirando s’iba, mirando mirando s’alguien la vía,
 8 vîra vîr un caballero que d’amor la pretendía.
 —Por Dios cho pido, caballero; por Dios y Santa María
 10 que me deixes ir con honra á esta santa romería.
 —Non t’hei deixar ir con honra ou te hei de quitar a vida.
 12 Comienzan de vuelta en vuelta, puñal d’oro lle caía;
 la niña como descreita⁶ muy de presa lo erguía:
 14 lo metió ‘n-el lado esquerdo y el derecho le salía.
 —Por Dios cho pido, romeira; por Dios y Santa María,
 16 que no digas en tu tierra ni tampoco en la mía
 que matach’ un caballero con las armas qu’ él traguía.
 18 —Hache ser tan secreto com’ a gaita en romería
 dind’a pirmeira ciudade hasta a última vila.

⁶ Mariño Ferro e L. Bernárdez (2002: 180) copiam mal *discreita*; igualmente não respeitam a contracção de 17a e transcrevem *q. matache un c.* Sobre a péssima qualidade e a falta total de critério científico do *Romanceiro en lingua galega* destes autores *vid.* a crítica de Forneiro (2002: 28) e a contrarréplica de Mariño e Bernárdez (2002^b: 28) onde amostram mais uma vez a súa incompetência no tema, além de evitar ou mascarar as críticas apontadas com um tom desrespeitoso.

Variantes AMP:

14a: l. m. n. l. ezquierdo

Variantes SEG:

Quintela do Pando (Viana)

O verso 2 está elidido.

4b: ô uso de Andalucía

5b: q. p. seda f.

8: v. v. u. cabaleiro q. d. a. l. requería

9: P. D. ch' o p. cabaleiro p. D. e S. M.

11b: que t. h. d. quitar a v.

12: Cabaleiro ô dar a volta p. d. ouro l. c.

13: A nena c. d. m. d. p. o e.

14: meteull'o do lado esquerdo y o corazón lle partira

15b: p. D. e S. M.

16: q. non d. na tua terra nin tampouco na miña

17a: q. m. u. cabaleiro cu'as a. q. e. t.

19: dende a p. c. h. na u. v.

A romeira II (b)

Palleirós (Puebla de Trives). Josefa Hervella.

Por aqueles campos verdes linda romeira camina⁷.

- 2 Saya leva sobre saya basquiña sobre basquiña;
 zapato leva picado ô uso d' Andalucía;
- 4 peinando vai su cabelo todo al doble le cubrira;
 media levaba encarnada que de lonxe relucía.
- 6 Mirando vai car' atrás, mirando s'alguen⁸ a vña,
 la vña un caballero de los que la pretendían.

⁷ *camiña* (Mariño & Bernárdez 2002: 180).

⁸ *alguen* (Mariño & Bernárdez 2002: 181).

- 8 —Por Dios cho pido, cabaleiro; por Dios e Santa María
que me deixes ir honrada a esa⁹ santa romería.
- 10 Anduveran volta á volta, puñal d'ouro lle caía.
La niña como discreta lo levantó par' arriba,
- 12 lo metió tras de la espalda y el corazón le partira.
—Por Dios cho pido, romeira; por Dios y Santa María
- 14 non-o digas 'n-a tua terra nin tampouco 'na miña
que matach' un cabaleiro co' as armas qu' él traguía.
- 16 —Non cho direi, cabaleiro, hasta 'n-a pirmeira vila,
qu' elí o hei de perdicar com'a gaita en romería.
- 18 En llegando á su casa ya su padre lo sabía
que matara un cabaleiro co' as armas qu' el traguía.
- 20 —¡Oh, banhaya a romeira y a madre que la parira,
que matóu un cabaleiro co'as armas qu' él traguía.

Variantes AMP:

Entre 9 e 10 figura o verso —Non t' hei deixar ir honrada, nin tampouco
â romería.

Variantes SEG:

San Esteban (Viana)

1b: l. r. camiña

4: p. v. s. cabeza c'un peine de prata fina

6: M. v. cara a. m. s'a. la vía

7b: d'aqueles qu'a p.

8a: P. D. ch'o p. c.

9b: a esta s. r.

Entre 9 e 10 figura o verso —Non t' hei deixar ir honrada, nin tampouco
â romería.

11: A meniña qu' é d. ergueu'o logo pra riba

12: meteull'o por entr'as costas y o c. lle p.

13: p. D. ch'o p. r. p. D. e S. M.

⁹ *esta* (Mariño & Bernárdez 2002: 181).

- 14: que n. d. na t. t. n. t. na m.
15a: q. matache u. c.
16a: N. c. hei dicir, caballero
17a: que ch'a de ser tan secreto
18: Cando chegou â s. c. xa seu pai o sabía
19: q. m. u. caballero
20: O. banhayas tu, romeira, y a. m. q. te p.
21: q. matache u. c.

A romeira III (c)

San Mamed (Viana del Bollo). Vicenta Ramona Fernández.

- Por aqueles campos verdes linda romeira camiña.
2 Saya leva sobre saya basquiña sobre basquiña;
zapato leva picado ô uso d'Andalucía;
4 peinando s'iba peinando con peines de prata fina.
Mirando se va, mirando, mirando s'alguién la vía:
6 Vírala un caballero; d'amores la pretendía.
La niña como discreita, dejó d' andar y corría;
8 el caballero d'a caballo muy pronto l' alcanzaría¹⁰.
—Dame tu cuerpo, la blanca; dame tu cuerpo, la niña;
10 dame tu cuerpo, la blanca, sino te quito la vida.
—Por Dios cho pido, caballero; por Dios e Santa María,
12 que me deixes ir con honra á esta santa romería.
—Dame tu cuerpo, la blanca, dame tu cuerpo, la niña;
14 dame tu cuerpo, la blanca, sino te quito la vida.
La niña con el temblor dijo que se lo daría;
16 levanta saya de grana, baja su blanca camisa,
se tiró 'n-un pozo hondo, donde la mar convertía
18 Todas las campanas tocan; todas tocan á alegría.
Unos dicen: ¿Qué será? y otros dicen: ¿qué sería?
20 Es la hija del caballero que para el cielo camina.

¹⁰ *la alcanzaría* (Mariño & Bernárdez 2002: 181).

A romeira ¹¹

- ^aPor aqueles campos verdes linda romeira camiña,
 2 ^bleva saya e sobresaia, basquiña sobre basquiña,
^azapato leva picado a estilo de Andalucía;
 4 ^bpeiteando vai seus cabelos ^acon peine de prata fina,
seus cabelos que de longos todo o alto lle cubrían.
 6 ^bVay mirando cara atráse mirando se alguén a vía;
^aben a viu un caballero que de amor a pretendía.
 8 ^cA nena como discrenta en vez de andare corría;
^ccabaleiro de a cabalo axiña a alcanzaría.
 10 *De lonxe lle dice “agarda, agarda, agarda, la niña.*
^cDame tu cuerpo, romera, sino te quito la vida”
 12 ^a—”Por Dios pido, caballero, por Dios e Santa María,
^aque me deixes ir con honra a esta santa romeiría.”
 14 ^a—”Ou te hey de quitar la honra ou te hey de quitar la vida”
^aComenzou de volta en volta; ^bpuñal de ouro lle caía;
 16 ^aben nena como discrenta desde o chan o recollía;
^bmeteullo pol’ as espaldas y o corazón lle partía.
 18 ^b—”Por Dios cho pido, romeira, por Dios e Santa María,
^aque non digas en tu tierra y ^bnin tampouco ^aen la mía,
 20 ^aque matache un caballero con las armas que él traguía”.
^b—”Non o direi, caballero, hasta na pirmeira vila,
 22 ^aque che a de ser un secreto como gaita en romeiría.”
^bEn achegando a sua casa xa seu pai ben o sabía,
 24 ^bque matara un caballero con as armas que el traguía.
^b—¡Oh, benhayas ti, romeira, benhayas ti, miña filla,
 26 ^bque matache un caballero con as armas que él traguía.

Cotarelo acompañava esta versión de *Una fatal ocasión* com a seguinte glosa:

Recogido de la tradición oral.

Poseemos cinco¹² versiones de este romance: las dos primeras se completan mutuamente; la tercera presenta bastantes castellanismos; la cuarta consigna

¹¹ Os autores do *Romanceiro en lingua galega* não repararam em nenhum momento que a versão de Armando Cotarelo estava elaborada com os textos de Hervella (Mariño & Bernárdez 2002: 176 e 180-182).

¹² Indicar três em lugar de cinco, bem como o ambíguo *recogido de la tradición oral* anterior, são marcas para distanciar-se das suas fontes.

algunos detalles obscenos innecesarios, y la quinta está contaminada por otro romance castellano moderno como se ve en el final: Dice hablando de la heroína:

*Se tiró nun pozo hondo donde la mar convertía.
Todas las campanas tocan, todas tocan a alegría;
unos dicen: “¿Qué será?” y otros dicen: “¿Qué sería?”
—Es la hija del caballero que para el cielo camina¹³.*

Con lo cual el argumento varía por completo.

Parece tratarse de un romance genuinamente gallego, parte de un ciclo que debió existir en torno a la romería de Santiago, y seguramente antiguo. Así lo indican toda su estructura, la repetición de ciertas frases y la acentuación paroxítona, para favorecer la cual acude a la adición de la e paragógica, sin necesidad, pues no afecta al asonante (verso 6)¹⁴.

El verso 5 parece interpolado¹⁵. La primera mitad del 8 semeja sustitución de otro perdido; nótese su repetición(sic) en el ¹⁶.

El 22 debe ser adición moderna. En los 12 y 18 contiene una frase imprecativa característica de los romances galaicos antiguos¹⁶ y los 20, 24 y 26 el pensamiento fundamental de la composición. Véanse en los 7, 10, 11, 14, 19 y 20 muestras de lo advertido acerca del uso del castellano por los personajes exóticos o empingorotados que figuran en los romances¹⁷.

A nuestro juicio pueden suprimirse los versos 3, 5¹⁸, 8, 9, 14 y 22 para la restauración del texto primitivo.

Em 1920, Armando Cotarelo inicia a série que prometera um ano antes e edita uma versão, publicada pela primeira vez em 1888 por António de la Iglesia, de *La mala*

¹³ Correspondem aos quatro versos finais da versão e de Hervella.

¹⁴ Ao nosso parecer, este e paragógico foi incluído pelo próprio Cotarelo a fim de ligar este romance com a lírica medieval.

¹⁵ É criação total de Armando Cotarelo.

¹⁶ Quais são esses romances galegos antigos? Cremos que aqui o nosso estudioso volta a tentar relacionar os romances com a lírica medieval, em concreto com as *Cantigas de Santa Maria* de Afonso X, obra a que já dedicara a sua atenção.

¹⁷ Cotarelo castelhaniza os versos das versões de Hervella com o fim de que o texto reflecta de maneira simplista a identificação das classes altas com o castelhano e as classes baixas com o galego. Esta descrição da diglossia da sociedade galega, que Cotarelo retratou nas suas obras teatrais bilingues, não se produz necessariamente nos romances da Galiza.

¹⁸ Inventado por Cotarelo. No entanto, não lhe parece que o verso 10, também composto por ele, pudesse ser prescindível.

suegra. Aqui, Cotarelo indica as referências bibliográficas do texto, uma versão que, como demonstrámos num trabalho anterior, não passava de ser um romance factício elaborado com materiais autênticos, que foram submetidos à manipulação (Forneiro 1999: 666).

Três anos depois, em 1923, Armando Cotarelo publica *Sinxebra*, peça teatral em língua galega, em que se incorporam versos de romances. Num dado momento, um grupo de rapazes canta o princípio do romance infantil *Hilo de oro*:

*De Francia vengo, señora, de comprar hilo francés,
en el camino me han dicho que lindas hijas tenéis* (Cotarelo 1923: 26)

Ambos os versos bem poderiam ser os da versão que Alfonso Hervella recolhera de Manuela Diéguez, em Santo Estevo (Viana do Bolo), com a correcção castelhanizadora de *han dicho* em lugar de *dijeron* do segundo verso, já que a única versão publicada deste tema, até esse momento, obtida por José Pérez Ballesteros, apresentava maiores diferenças¹⁹. Umhas páginas depois aparece o verso —*Me casó mi madre chiquita y bonita*, do também romance infantil *Me casó mi madre* tema só recolhido na tradição oral da província de Lugo, portanto, talvez obtido pelo próprio Armando Cotarelo. Também na mesma página se encontram os seguintes versos, um arremedo do estilo do romanceiro tradicional, pois não correspondem a nenhum tema testemunhado na Galiza:

—*¿Por qué lloras, bella niña? —¿Por qué tengo de llorar?*
Mi mamáta se ha muerto, mi padre se va a casar
y a mí y a mis hermanitos madrastra nos quiere dar.
Por eso lloro, señora, ¡bien tengo por qué llorar! (Cotarelo 1923: 32)

Um ano depois, Cotarelo inclui noutra peça teatral, *Lubicán*, um romance, desta vez inteiro, onde nas páginas 11-13 imita o estilo do romanceiro tradicional. Antes da recitação deste romance a sua transmissora, a tia Miñota, menciona que conhece, entre outros, *Delgadiña*, *San Xoán*, *A romeira*, *Frol da auga*, *Tres comadres*²⁰, todos estes temas estavam presentes no trabalho de Alfonso Hervella.

¹⁹ De Francia vengo, señora, de un pulido portugués,
en el camino me ha dicho que lindas hijas tenéis
(Pérez Ballesteros 1884: 136)

²⁰ O título *Tres comadres* não aparece na lista de temas de “Romancero popular gallego”. *San Xoán e Frol da auga*, são, provavelmente, o mesmo tema com títulos diferentes.

Em 1925, Armando Cotarelo publica na revista *Nós* o artigo “Un romance galego”, trabalho lido no Seminário de Estudos Galegos em 15/10/1924, nele, como já indicámos antes, editava uma versão do romance bíblico de *Tamar*²¹, que não passava de ser um texto factício elaborado com duas versões recolhidas por Hervella. Eis os textos:

El incestuoso I (a)

Espino (La Vega). José Nieto Carracedo.

- Tres hijos tenía el rey, todos tres com'una grana,
2 el más chiquitito d'elles se namoró d'una hermana.
Un día por la gozar s'hallaba malito en cama.
4 Bajara su padre á verlo d'altas torres dond' estaba.
—¿Qué tienes tú, hijo mío? ¿qué es lo qu' á ti te pasa?
6 —Calientura tengo, padre, calientura que no m'arranca.
—¿Qué tomarás, hijo mío, qué tomarás que te 'o traiga?
8 —Un guisado, padre mío, que me lo hiciera mi hermana.
—Que lo haga, hijo mío; que te lo haga y que lo traiga.
10 —¿Qué tienes tú, hermanito, qu' estás tu malito en cama?
—Todo esto, mi hermana, tus amores me lo causan.
12 —Mis amores, hermanito, á ti no t'importan nada.
La cogió por la cintura y en la cama la echara;
14 voces daba la doncella á su padre que le valga.
—Cállate tu, hermanita, tod'esto no importa nada:
16 tenemos un tío en Roma que todo lo dispensaba.

Variantes AMP:

6b: c. q. n. m'arrama

Variantes SEG:

Palleirós (Puebla de Trives)

2a: e. m. pequenito d' eles

²¹ Esta versão sob o título de *Torpinos* (*Versión recogida por Dn. Armando Cotarelo*) foi incorporada à coleção (pp. 30-32) que o advogado salmantino Alejo Hernández obteve na cidade da Corunha e na sua periferia, e no norte de Lugo, durante os anos 1924-1925 e que hoje faz parte do *Archivo Menéndez Pidal*.

6: Calentura t. p. calentura q. n. arrama

8a: U. g. mi padre

9b: q. t. l. h. y l. t.

15b: t. e. n. val n.

16b: q. t. esto d.

El incestuoso II (b)

Palleirós (Puebla de Trives), Josefa Hervella.

- Tres hijos tenía el rey, todos tres como una grana;
2 el más chiquitito d'eles se namoró d'una hermana;
para la vencer más presto s'hallara muy malo en cama.
4 Un domingo á la tarde vino por allí su madre.
-¿Qué comerás, fillo meu? ¿qué comerás que ch' eu traiga?
6 -La pechuga d'una pava si me la guisa mi hermana.
-Eu lle direi que ch'a guise; eu lle direi que ch'a traiga.
8 -Si viene que venga sola, que no venga acompañada,
que la bulla de la gente á los enfermos enfada.
.....
10 Le cogiera la pechuga, la tirara por la sala;
la cogió por los cabellos, por la sala l'arrastrara,
12 tanta burla fizo d'ela qu' hasta le cuspió en la cara.
Se saíra car' afora la probe desconsolada.
14 -Justicia veña do Ceyo, qu'a da terra non val nada,
pra matar o fillo do Rey que mi casa estropeara.

Variantes AMP:

1b: com'una grana

14a: Xusticia v. d. c.

Variantes SEG:

San Esteban (Viana del Bollo)

2a: e. m. pequenito d.

4b: su madre ond' el llegaba

6a: l. p. de una p.

9a: porque el ruido d. l. g.

10: La cogió por la cintura y en la cama la echara

-¡Válgame pronto, mi padre, Válgame la Virgen Santa!

12: -t. b. hizo d'e.

13: Saliérase para fuera l. pobre d.

15: p. matar o meu irmau q. mi honra me quitara

Versão de Armando Cotarelo

- ^{ab}Tres hijos *tiña* este rey, todos tres com' una grana;
 2 ^bel más chiquito *de eles* se namoró de una *hirmana*;
^bpara la vencer mas presto se hallara muy malo en cama.
 4 ^aBajara su padre a verlo *de altas torres onde* estaba.
^a—¿Qué tienes tú, hijo mío, qué *es eso* que a te ti pasa?
 6 ^a—Calientura tengo, padre, calientura que no arranca.
^b—¿Qué comerás, fillo meu? ¿qué comerás que *eu che traia*?
 8 ^a—Un guisado, padre mío, ^bsi me lo guisa mi hermana.
^b—Eu *llo* direi que *cho* guise, eu *llo* direi que *cho traia*.
 10 ^b—Si viene que venga sola, que no venga acompañada,
^bque la bulla de la *giente* a los enfermos enfada.
 12 ^a—¿Qué tienes tú, *mi* hermanito, qu' estás *tan* malito en cama?
^a—Todo esto, hermana *miña*, tus amores me lo causan.
 14 ^a—Mis amores, hermanito, á ti no te importan nada.
^aLa cogió por la cintura y en la cama la *tumbara* ;
 16 ^avoces daba la doncella a su padre que *la* valga.
^a—Cállate tu *mi* hermanita, todo esto no importa nada:
 18 ^atenemos un tío en Roma que todo lo dispensaba.
^bTanta burla fizo *de ela* que hasta le cuspió en la cara.
 20 ^bSe saíra *cara* afora *berrando* desconsolada.
^b—*Xusticia* veña do ceio, *que a* da terra non val nada,
 22 ^bpra matar *fillo* do Rei que *a min me deixa bulrada*.

A 7 de Abril de 1928, Armando Cotarelo dá a lume *O Pranto da Virxen* numa página literária de *El Compostelano*, onde este romance tradicional estaba acompañado de poemas de Manuel de Sandoval, Teixeira de Pascoaes, Rafael Sanz de Diego, Luis Contreras e Guillermo F. Shaw²². O título com que Cotarelo identificava este tema era a tradución para galego de *El llanto de la Virgen (Por las almenas del cielo no, IGR)*, como assim intitulara Alfonso Hervella as dúas versións que obtivera na súa vila natal. Armando Cotarelo, aquí como nos dous casos anteriores, fabricou o seu texto misturando, ao mesmo tempo que alterando, as versións da colección de Hervella.

El llanto de la Virgen I (a)

San Esteban (Viana del Bollo). Aquilina Carballo.

Por los caminos del Cielo s'apasea una doncella,
 2 toda vestida de blanco, cubierta de gracia llena.
 Preguntó San Juan a Cristo: “-¿Quién es aquella doncella?
 4 Y Cristo le respondió:
 —Aquella es Nuestra Madre, Nuestra Madre verdadera.
 6 —Pues que es Nuestra Madre, adoremos todos ‘n-ella;
 Los angelitos del cielo y los hombres de la tierra
 8 han visto llorar la Virgen debajo d’una olivera.
 —¿Por qué llora usted, Señora? ¿por qué llora usted, doncella?
 10 —Lloro por una noticia que me trougo Madalena:
 que un hijo que yo tenía, que sin dolor lo pariera,
 12 me lo están crucificando en una cruz de madera.
 Un poquito más alante hay un sepulcro de piedra,
 14 donde murió Jesucristo, Salvador de Cielo y tierra;
 y un poquito más alante hay una cruz de madera,
 16 donde murió Jesucristo, Salvador de Cielo y tierra;
 no murió por sus pecados, que nunca los cometiera:
 18 murió por los pecadores por nos dar la Gloria eterna.

²² A fotocópia desta página de *El Compostelano* (assim como outra da primeira página deste jornal de 31 de Março de 1927 em que Alfonso Hervella Courel publicava una versión de *La infantina*.) foi-me oferecida por Armando Requeixo, a quem agradeço aquí a súa amabilidade.

Variantes SEG:

1b: se apasea u. d.

4a: Elidido

5: P. qu'es N. M. todos adoremos 'n ella

El llanto de la Virgen II (b)

San Mamed (Viana del Bollo). José Casal.

- Por los caminos del Cielo s'apasea una doncella,
2 toda vestida de sangre, cubierta de gracia llena.
Preguntó San Juan á Cristo: ¿Qué doncella es aquella?
4 -Pues aquella es Nuestra Madre, Nuestra Madre verdadera.
—Pues ya qu'es Nuestra Madre, todos adoremos 'n-ella.
6 Los angelitos del Cielo y los hombres de la tierra.
han visto llorar la Virgen por debajo d'una olivera.
8 —¿Por qué llora usted, Señora? ¿Por qué llora usted, doncella?
Lloro por una noticia que me trajo Madalena;
10 por un hijo que 'o he tenido que sin dolor lo pariera,
que está muerto 'n-una cruz, 'n una cruz de madera.
12 —Si me lo bajan devotos muy contenta m' estuviera;
así con fuerzas me hallara que yo misma lo cogiera.
14 Allí 'n-el monte Calvario hay una cruz de madera,
ond' está el Salvador, Salvador de Cielo y tierra;
16 no murió por sus pecados, que nunca ningun tuviera:
murió por los pecadores por nos dar la Gloria eterna.

O Pranto da Virxen

“Recogido de la tradición oral, por Armando Cotarelo”.

- ^{ab}*Polos camiños* del Cielo se *paseia* una doncella,
2 ^atoda vestida de *branco*, *coberta de gracia chea*.
^bPreguntó San Juan a Cristo: “—¿Qué doncella *será* aquella?
4 ^b—*Pos* aquella es nuestra Madre, ^atodos adoremos nella.

- ^aLos angelitos del Cielo y los hombres de la tierra
 6 ^ahan visto *chorar a Virxen debaixo de unha oliveira.*
^{ab}—¿Por qué *chora* usted, señora? ¿por qué *chora*, usted, doncella?
 8 ^a—*Choro por unha* noticia que me trougo Madalena:
^aque un hijo que yo tenía, que sin dolor lo pariera
 10 ^ame lo están crucificando en una cruz de madera.
^bSi me lo bajan devotos muy contenta me estuviera
 12 ^by *si* con fuerzas me hallara yo mesmo lo recogiera
^a*que* un poquito más alante hay un sepulcro de piedra.
 14 ^bAllí *en el* monte Calvario *está* una cruz de madera
^adonde murió Jesucristo Salvador de cielo y tierra.
 16 ^aNo murió por sus pecados que nunca los cometiera,
^{ab}murió por los pecadores por nos dar la gloria eterna.

Ignoramos se Cotarelo publicou na imprensa mais textos procedentes do acervo coligido por Alfonso Hervella²³, mas os três textos aquí analisados manifestam tanto a ultrapassada concepção do romanceiro tradicional de Armando Cotarelo, quanto a sua desonesta apropriação de materiais alheios. Os seus preconceitos clasistas e letrados faziam com que considerasse que a tradição oral estropiava os textos que adoptava como próprios: “[...] los romances se hallan hoy atrozmente desfigurados, según ocurre con cuanto no se fija en la escritura” (Cotarelo, 1919: 35), “[...] este romance, como tantos otros, al transmitirse por tradición oral [...] ha sido notablemente desfigurado con repeticiones, interpolaciones, cambios de palabras y de tiempos de algunos verbos, supresiones, lagunas, etc. (Cotarelo 1920: 277), “[...] o romance amostra ser moi antigo, inda que a forma en que o puiden recoller háchase estragada de tanto peneirar de boca en boca” (Cotarelo, 1925: 2)

A consideração de a tradição oral corromper os romances que acolhia explica que a manipulação e o retoque dos textos por parte dos colectores e editores fosse prática comum na segunda década do século XIX e, nalguns casos, até nos primeiros anos do século XX. Todavía, o respeito pelos textos orais foi exigido pelos estudiosos positivistas das últimas décadas do século XIX. Assim, por exemplo, na Bretanha da década de 60 foi posta em causa a autenticidade do *Barzaz Breizh* (1839), colecção de baladas bretãs reunidas por Hersart de la Villemarqué (Postic 1997: 97-123). Por sua vez, o folclorista e dialectólogo sueco A. W. Munthe censurou, em 1888, que Juan Menéndez Pidal criasse versões factícias com os roman-

²³ As bibliografias da sua obra não indicam de que Armando Cotarelo editasse mais romances (Barreiro 1984: 3-4 e Varela 1984: 135-148).

ces tradicionais na sua *Colección de los viejos romances asturianos* (1885) (Cid, 1992: 45), a primeira colección moderna importante em língua castelhana. O próprio Alfonso Hervella também foi um colector positivista da tradição oral galega como se pode comprovar no último parágrafo do prefácio dos seus *Romances populares gallegos* :

No terminaremos este ya enfadoso prólogo sin hacer constar que todas las versiones incluidas en esta colección son fidelísimas copias de las narraciones orales, que personalmente recogimos de labios de sus recitadores; con lo dicho queda asegurado que no hemos incurrido en el imperdonable vicio del amaño y el retoque. No hemos querido imitar a los románticos en lo de “dorar el oro y platear la plata de la poesía popular” (Hervella, 1909: XXIV).

Não deixa de resultar curioso que um advogado interessado pelo romanceliro da sua terra fosse um editor de romances mais científico²⁴ que Armando Cotarelo Valledor, o principal filólogo galeguista das décadas de 20 e 30 do século passado. Se Cotarelo recorreu ao “vicio del amaño y del retoque” sobre os textos de Hervella, não só foi por julgar a literatura tradicional como um saber inferior à literatura considerada culta, pois também pretendia apropriar-se de materiais que lhe permitissem ser considerado o maior especialista nos romances galegos. Uma vez que Manuel Murguía se ocupara e sublinhara a importância deste género da literatura tradicional na cultura do país, e que Víctor Said Armesto (o homem de letras mais prestigioso na Galiza de inícios do século XX) recolhera entre 1900 e 1910 um excelente e riquíssimo romanceliro nas quatro províncias galegas²⁵, Armando Cotarelo sentiu-se obrigado a tratar do romance de tradição oral no seu afã por ser o patriarca da cultura na Galiza daquela época.

No entanto, para além desta razão, cremos que o interesse de Cotarelo pela literatura popular também respondia à sua visão política, social e cultural da Galiza e da Espanha. Se repararmos nos assuntos dos três romances aqui analisados, podemos ver que todos têm uma relação mais ou menos directa com a religião católica. Deixando de lado o tema de *O Pranto da Virxen*, pelo óbvio do seu assunto e porque Cotarelo não o acompanhou duma glosa, os romances de *A romeira* e de *Tamar* fornecem-nos reveladoras informações sobre a ideologia do nosso autor.

²⁴ De qualquer modo, Alfonso Hervella nem sempre respondeu à manifestação positivista do seu prefácio; os seus preconceitos morais, devidos ao seu conservadorismo católico, explicam que nalguma ocasião suprimisse versos ou vocábulos dos romances *pudoris causa*, segundo as suas palavras.

²⁵ Cotarelo Valledor manifestou sobre a coleção de Said: “Hace años que vimos en Madrid esta colección en manos de su autor y nos pareció muy notable. Ignoramos su paradero. Algunos de los romances que contenía han sido publicados en diferentes ocasiones [...]” (Cotarelo, 1919: 33).

No caso de *A romeira* Armando Cotarelo relacionava este romance, como já vimos, com o caminho de Santiago, baseando-se no facto de a protagonista ser uma romeira; porém, o tema carece duma qualidade especificamente galega, dado que é uma balada difundida em toda a Europa. A ligação deste romance com a matéria jacobea, unia a reivindicação do seu carácter galego ao mesmo tempo que o seu valor religioso. Se o galeguismo progressista opôs desde a Geração *Nós* até aos nossos dias um Apóstolo Santiago europeísta e universalista ao tradicional Santiago matornouros do espanholismo mais castiço e integrista (Barros 1994, 253), isto não deve levar a pensar que Cotarelo, pelas suas actividades galeguistas nas primeiras décadas do século XX, não participasse da concepção do mito jacobeu própria do tradicionalismo espanhol.

Em nossa opinião, o abuso da identificação na historiografia galeguista da pós-guerra do nacionalismo com a história da Galiza (Ventura 2000: 95 e Villares 1984: 164) e do uso da língua galega com valores progressistas e com atitudes “arredistas” ou anti-espanholas têm impedido a correcta compreensão de certas tendências no galeguismo ao longo da história, bem como dos percursos políticos e intelectuais dalguns vultos galeguistas. Hoje o nacionalismo galego costuma apagar no relato historiográfico tudo aquilo que não coincide com os seus actuais valores e, destarte, esquece ou minimiza a existência dum galeguismo conservador²⁶, surpreende-se perante a afervorada adesão dos regionalistas à causa espanhola durante a crise de 1898 (Casares 1998: 5-23), ou deixa de lado factos importantes numa biografia humana e intelectual como o uso exclusivo do castelhano por parte de Manuel Murguia e Rosalia de Castro no seu quotidiano (Durán 1977: 290-291).

O estudioso catalão Joaquim Ventura contestou recentemente no seu *O nacionalismo cármico de Vicente Risco* (2000) a visão canónica do nacionalismo galego actual sobre a evolução ideológica deste vulto da cultura galega: de principal teorizador do galeguismo antes de 1936 a fascista converso a partir dessa data. No entanto, tal percurso ideológico, como muito bem demonstra Ventura, nunca teve lugar. Vicente Risco foi sempre o mesmo antes, durante e depois da sua experiência galeguista: um teósofo reaccionário temeroso das inovações materiais e intelectuais da sociedade contemporânea ocidental (Ventura 2000: 93-94; 118-122). Nem foi progressista por aderir ao nacionalismo galego da pré-guerra, nem se tornou fascista a partir do golpe de estado do General Franco (Ventura 2000: 131-141). Neste sentido, o caso de Vicente Risco bem poderia ser o de Cotarelo Valledor.

²⁶ Se bem que seja indubitável que a tradição predominante no galeguismo contemporâneo tem sido a progressista, graças ao liberalismo de Murguia, Vilar Ponte e Castelao e ao desinteresse da Igreja galega pela língua e a cultura autóctones.

O facto de Armando Cotarelo ter dinamizado a vida universitária, cultural e social na conservadora Compostela de principios do século XX, ou de se comprometer com o galeguismo (presidindo o Seminario de Estudos Galegos e empregando a lingua galega nos usos coloquiais (Filgueira 1984: 15 e Carvalho Calero 1981: 111) e literários), não devem levar a engano: a militância galeguista de Armando Cotarelo durante a súa estada na Galiza não era qualquer amostra de um progressismo que trairia a partir de Julho de 1936. Formado intelectualmente dentro das coordenadas ideológicas do integrismo hispânico (lembramos que era fillo de Emilio Cotarelo Mori, um dos seus representantes mais notáveis junto a Marcelino Menéndez y Pelayo), Cotarelo foi sempre coerente com o seu conservadorismo inicial, fosse duma perspectiva espanhola ou galega. Daí que a súa “aventura” galeguista nunca contradísse o seu espanholismo prévio e posterior, uma vez que o seu galeguismo nunca fora anti-espanhol nem o seu uso da lingua galega acarretara grandes mudanças na súa praxe intelectual²⁷. No editorial do primeiro número da revista *Utreya* (1919) que ele dirigia (portanto, poucos anos antes de empregar o galego como lingua literária), podiam ler-se estas significativas palabras sobre a relación entre a Galiza e a Espanha e sobre a lingua galega:

Somos españolistas porque somos españoles y porque sabemos que cuanto se haga por España tanto se hará por Galicia, y somos regionalistas porque somos gallegos y por la convicción en que estamos a laborar por la gran patria española.

Por eso nosotros, aunque sentimos hondamente el habla gallega y anhelamos fomentar su empleo en determinados géneros de la literatura regional, no somos partidarios de su cooficialidad con el castellano, pues todo gran pueblo necesita de una gran lengua común para desarrollarse y consolidar su personalidad, y sería injusto negar tal prerrogativa a la lengua de Castilla, que tiene tan gloriosa historia y se habla por cien millones de habitantes en veinte florecientes estados (Cotarelo 1919: 2).

O uso coloquial da lingua galega por Armando Cotarelo, de acordo com as novas prácticas linguísticas dos galeguistas do século XX²⁸ e a súa produción literária bilin-

²⁷ De facto, parece que a súa figura não suscitou adesões unánimes dentro do movemento galeguista: “Queda tamén sen aclarar cal é a súa verdadeira posición política na paisaxe política da época. A súa figura provoca nos galeguistas posicións encontradas, desde a gabanza ó ataque, que ben pagaría a pena pescudar. Sirva como mostra disto un significativo artigo de Villar Ponte no que afirma que “Nuestro ilustre amigo Don Armando Cotarelo no ha sido nunca capaz de entender a Risco ni entendernos a nosotros cuando hablamos de nacionalismo” (Boullón-Tato 1991: 40-41).

²⁸ Ao que parece, os escritores oitocentistas em lingua galega nunca ou raramente empregavam a fala autóctone nos usos orais, segundo os testemunhos de Emilia Pardo Bazán (1984: 296) e de José de la Hermida, primo de Rosalía de Castro (Durán 1977: 290-291).

gue ou em língua galega (peças teatrais de assunto histórico e sobretudo rural²⁹), bem como alguns trabalhos científicos na língua do país (*Cancioneiro da agulla*, *Un arcebispo de Santiago descoñecido*, *Etimoloxías populares*, etc.), conviveram com o uso do castelhano nos seus romances e na maior parte da sua obra científica. Inclusive, como já vimos anteriormente, Cotarelo também praticou o bilinguismo quando manipulou o romanceiro tradicional³⁰, à diferença de editores de romances do século XIX, como Murguía e Antonio de La Iglesia, que traduziam para galego tudo o que aparecesse em castelhano nos romances recolhidos do acervo popular.

Os galeguistas conservadores das primeiras décadas do século XX, Cotarelo, Vicente Risco, Filgueira Valverde ou Noriega Varela, não traíram a sua Galiza ideal por aderir ao General Franco a partir de 1936. Todos eles foram, de maneira quase inevitável, franquistas, porque perante tudo eram católicos integristas e daí que o seu galeguismo não passasse de ser a forma que numa dada altura certo integrismo espanhol adoptara para, desde *la región*, resistir às “devastações” da modernidade: o desenvolvimento tecnológico, a urbanização, o laicismo e as ideologias de esquerda.

Por outro lado, o galeguismo de Cotarelo era uma continuação do nacionalismo endocêntrico que caracterizou as aristocracias do Antigo Regime, que reivindicavam para as suas respectivas comunidades ser o berço das Espanhas. Uma vez que os bascos foram durante esse período histórico os depositários do essencialismo espanhol por sua limpeza de sangue semita³¹, para Cotarelo era inaceitável que na Espanha Contemporânea um povo do norte perdesse semelhante privilégio em favor dos andaluzes³², “os semitas espanhóis” por excelência. Em nossa opinião,

²⁹ São de interesse as seguintes palavras de Ramon Reimunde sobre *Trebón*, que bem se poderiam aplicar ao resto da produção literária de Armando Cotarelo: “Mais nom hai na sua obra *Trebón* nengum elemento de propaganda galeguista nem de sátira contra a desgaleguizaçom como outras do seu tempo (*Mal de moitos*, 1915), nem dado algum para conhecer a situaçom sócio-económica e cultural do campo em Galiza a começos do século XX” (Cotarelo, *T*: 37).

³⁰ Também os dois grandes colectores de romances de princípios do século XX, Alfonso Hervella e Víctor Saíd, retocaram os romances que obtiveram da tradição oral empregando o galego e o castelhano.

³¹ Conforme a antropóloga flamenga Christiane Stallaert “el nacionalismo vasco, nacido por la competición por la limpieza de sangre, es el fruto más puro que ha producido el etnicismo español” (Álvarez Chillida 2002: 233).

³² A identificação da Andaluzia com a Espanha foi criada pelo Romantismo francês e inglês, após a derrota do exército de Napoleão na Península Ibérica. Sirva, de exemplo, o seguinte trecho de Félix de Azúa ao respeito: “Aquel viajero que a principios de siglo podía escribir impunemente: ‘Al fin me encuentro en Irún, es como Constantinopla’, haría el ridículo si en el período post-napoleónico no fuera más al sur en busca de novedades. Así, poco a poco, por acumulación informativa y a medida que Vasconia, Cataluña o Valencia van viéndose como tierras enteramente normales, como cosa común y europea, aunque más pobre y aburrida; a medida que van dejándose de ver espectaculares primitivos en

depois de o recém nascido nacionalismo basco rejeitar a espanholidade do seu país quando o liberalismo afundou a identidade castiça em Espanha (Álvarez 2002: 235), e dado que os andaluzes pelo seu passado muçulmano não eram, em termos ráricos, “de confiança”, Cotarelo propunha que a Galiza preservasse na Espanha do século XX, as essências que caracterizaram a identidade espanhola durante a Idade Moderna: a limpeza de sangue e a religião católica.

Segundo Hannah Arendt, não existe uma relação filial entre o sentimento anti-judeu próprio dos confrontamentos entre diferentes credos religiosos antes da Idade Contemporânea e o anti-semitismo, uma ideologia secular oitocentista que considerava os judeus na altura da crise do Estado-Nação europeu como um grupo *anti-nacional* pelo seu carácter internacional (Arendt 1968: 25). Porém, para outros estudiosos, o anti-semitismo moderno bebe do velho anti-judaísmo (Karady 2000: 210 e Álvarez Chillida 2002: 215). Seja como for, de acordo com Jon Juaristi, naqueles estados nacionais europeus que careciam duma minoria judia, como era o caso espanhol, procurou-se um grupo que ocupasse “o gueto vazio”. Assim, para o tradicionalismo hispânico os liberais espanhóis substituíram os judeus, uma vez quebrado, com a Constituição de Cádiz (1812), o pacto nacional da Época Moderna, baseado na unidade católica do povo espanhol e no seu correlato negativo, o anti-judaísmo (Juaristi 1994: 121-122)³³. Conforme com isto tudo, parece claro que a assunção por Armando Cotarelo do anti-judaísmo do Antigo Regime tinha como objectivo neutralizar os liberais, a nova anti-Espanha que substituíra a velha anti-Espanha árabe e judia. Isto explica o interesse pessoal e “regional” de Cotarelo quando reivindicava para a Galiza um papel de relevo na Espanha do século XX, pois o território galaico, livre de influências semíticas e caracterizado pelo seu secular isolamento, podia ser um firme baluarte na defesa dos “antergos” valores hispânicos face às dissolventes inovações intelectuais que acarretava a modernidade liberal.

Esta defesa do essencialismo rárico e religioso espanhol por Cotarelo patenteia-se em obras como *Fray Diego de Deza* (1902) e *Límite septentrional de la conquista sarracena en España* (1921). No primeiro destes livros, Armando Cotarelo ocupa-

Barcelona o en Alicante, Andalucía se dibuja como la última esperanza del romanticismo, como la verdadera patria. Y así, poco a poco, Andalucía se convierte en la super-España y los andaluces en super-españoles: hombres arcaicos de vida sexual libre (especialmente las mujeres por chocante que parezca), de extraordinarias aptitudes artísticas, ignorantes de su propia valía y viviendo en un paisaje salvaje de magníficos contrastes” (Azúa 1996: 165).

³³ Os nacionalismos periféricos também julgaram que o semita, ou o considerado como tal, era um perigo para a coesão nacional. No nacionalismo galego já foi repetidamente assinalado o caso de Vicente Risco e há pouco foi apontado o anti-semitismo dalguns vultos do nacionalismo catalão (Marfany 1995: 197-201; Álvarez 2002: 240-249) e do nacionalismo basco (Juaristi (1994: 115-136; Karady 2000: 226; Álvarez 2002: 229-240).

va-se de retratar a figura deste samorano, nascido em 1422 e de apelido galego (Cotarelo 1902: 18), que, além de mestre do príncipe Don Juan (filho dos Reis Católicos), foi catedrático em Salamanca, Bispo de Samora, Jaém e Palência, arcebispo de Sevilha e inquisidor geral. Segundo Cotarelo, não foi tão duro como o famoso Torquemada e

mostró siempre particular enemiga a los moriscos y cristianos nuevos procedentes de judíos; contra estos especialmente excitó a los Reyes Católicos, con motivo de haber venido a España varios distintos de los expulsados en 1492, y obtuvo Real Cédula despachada en 5 de Octubre de 1499, para que se extendiesen a ellos las providencias acordadas contra los otros.

Quiso también persuadir a los monarcas de que convenía poner Inquisición en Granada, pues los moriscos recién conversos abusaban de su falta y tornaban al mahometismo (Cotarelo 1902: 145)³⁴.

Em *Límite septentrional de la conquista sarracena en España*, Armando Cotarelo seguia o modelo do seu admirado Francisco Codera (de quem publicou uma biografia em 1898) que em 1879 tracejara o limite da invasão sarracena na Catalunha, Aragão e Navarra. Cotarelo concluía na sua obra:

De manera que tratando de señalar el límite superior de la Conquista sarracena en España sin vacilar nos inclinamos a establecerla en la línea formada por la gran cordillera Cantábrica hasta la frontera gallega y en esta comarca por el río Miño hasta su desembocadura en el Océano Atlántico (Cotarelo 1921: 42).

A intenção anti-semita de ambos os trabalhos é inequívoca, e no segundo caso podemos supor que Cotarelo Valledor estava interessado em sublinhar que a sua terra natal nunca fora invadida por povos africanos. Este anti-semitismo também se pode encontrar no artigo “Un romance galego”, concretamente na glosa em que Cotarelo comenta a versão do tema bíblico de *Tamar*, onde manifesta, de maneira muito explícita, a sua aversão por judeus e mussulmanos, bem como a sua vontade por vincular a Galiza com Castela, o eixo de Espanha:

A importancia diste eixemplar rube de súpeto ô considerar que Menéndez Pelayo esquirbiu referíndose a certa mostra romancesca andaluza do mesmo asunto hachada en Osuna por Rodríguez Marín [...] “es el único romance popular que conozco sobre asuntos del Testamento Viejo” [...]. Agora ben. Si

³⁴ “[...] Deza (a pesar de ser él mismo un marrano, de linaje converso por vía de hembra) aspiraba a una limpieza étnica de España mediante la desaparición (vía inquisitorial) de todos los cristianos nuevos procedentes de judíos o moriscos” (Catalán 1998: 100).

comparamos o romance de Osuna co noso, temos que lle dare a iste grande superioridade. É mais longo, mais compreto, mais movido, mais dramático e segurísimamente moi anterior. O andalus parece obra de algún xudeu ou morisco, según acara a anteposición do artello al ô nome de Tamar. A substitución de Amnón por Tarquino, o forzador de Lucrecia é un caso ben craro de iso que se chama contaminación entre dous temas, un de Ourente e outro de Oucidente³⁵.

O noso poeta non fixo nada d'esto. Con soilo o recordo de aquela bárbara páxina do Libro de Reises, teceu unha refenda³⁶ bárbara tamén culpa do asunto, pro mais xêral, mais human e millor condida e desempeñada. Eiquí non se ve rastro algún de arabismo ou xudaísmo; por isto quero creer que o romance non nasceu en chola semítica, senón que de Castela veu a Galiza onde os nosos paisans o aniparon garimosos e de Castela baixou a Andalucía onde ise mouro ou xudeu estragou nél de moi mal xorne (Cotarelo 1925: 3)

Em definitivo, no labor de Armando Cotarelo como editor e investigador da tradición romancística da Galiza, encontramos diversas chaves para melhor comprendermos a súa personalidade e a súa obra, apesar de o estudo do romancero ter ocupado un lugar secundário dentro da súa produción intelectual. A atención do erudito da Veiga do Eu pelo romancero, obedecía, por um lado, à súa ansia de protagonismo no mundo cultural galego de inícios do século XX, uma vez que o romance foi um tema a que Manuel Murguía dedicara grande atención, e porque na primeira década do século XX se reuniram na Galiza duas excelentes colectâneas romancísticas: as organizadas por Said Armesto e por Hervella Courel. Portanto, Cotarelo não podia pôr de parte um tema tão vivo na súa época e daí que chegasse, desonestamente, a tirar partido dos materiais coligidos por Hervella a fim de tornar-se num dos maiores especialistas galegos na matéria. Além disto, Armando

³⁵ Cotarelo copia aquí palabras de Menéndez Pelayo: “Puede ser obra de algún judío o morisco, como parece indicarlo la anteposición del artículo Al al nombre de Tamar. La sustitución de Amón por *Tarquino* o Tarquino (¿el forzador de la romana Lucrecia?) es un caso de contaminación muy singular entre dos temas poéticos: uno de Oriente y otro de Occidente” (Menéndez Pelayo 1900: 197).

³⁶ Esta “refenda” é uma má lectura. Na página VII do prefácio de Alfonso Hervella na cópia do seu romancero que enviou a Ramón Menéndez Pidal expõe: “[...] los romances genuinamente castellanos al ser interpretados por el pueblo gallego se hagan en cierto modo regionales e indígenas. Este carácter es más marcado, más propio e más exclusivo en aquellas cantigas⁽¹⁾ que se han inspirado en la historia local [...] (1) Así se llaman en la región los romances e también historias, cantares, referidas”. Este *referidas* na cópia do romancero de Hervella depositada no Museu de Ponte Vedra (p. VIII) lê-se como *refendas*, e Cotarelo na súa peza teatral *Lubicán*, também usa esta palabra: Mozas: ¡A refenda, a refenda! T. Miñota: Pois a tergo^(sic) escollede: *Delgadiña, San Xoán, A romeira, Frol d'auga, Tres comadres...*” (Cotarelo 1924: 11). Esta estranha palabra foi incorporada por Eladio Rodríguez no seu dicionário e na antoloxía desta obra lexicográfica editada por *La Voz de Galicia* no Día das Letras Galegas de 2001 é uma das palabras seleccionadas: “*Refendas*: Historias, contos, xeralmente de meigas e pantasmás” (Rodríguez 2001: 29).

Cotarelo serviu-se do romanceiro para reivindicar o papel da Galiza na Espanha do seu tempo. Deste jeito, Cotarelo asumía o castelhanismo originário e lingüístico do romanceiro tradicional porque (à diferença da maioría dos galeguistas desde o século XIX) pretendía colocar a Galiza dentro do concerto das Espanhas como o lugar que, para além de ter preservado o melhor de Castela, como o romanceiro, era um repositório do mundo castiço español, limpo das influências semitas de que tanto acusavam os andaluzes, “os mais espanhóis dos espanhóis” na Idade Contemporânea.

Referências bibliográficas

- Alvarez Chillida, G. (2002): *El antisemitismo en España. La imagen del judío (1812-2002)* (Madrid: Marcial Pons).
- Armando Cotarelo, T = Reimunde R. (ed.) (1984): Armando Cotarelo, *Trebón* (s. l.: AGAL).
- Azúa, F. de (1996) = González Sáinz, J. A. (ed.): Azúa, F., “El mito de la Andalucía romántica” (1981), *El aprendizaje de la decepción*, (Barcelona: Anagrama).
- Barreiro Fernández, X. R. (1984): “Cotarelo e a Universidade de Santiago”, *Revista Monográfica de Cultura*, 1: 5-10 (A Coruña: A. C. “O Facho”).
- Barros, C. (1994): “Mitos de la historiografía galleguista”, *Manuscripts*, 12: 245-266.
- Boullón Agrelo, A. I. e Tato Plaza, F. R. (1991): “A cátedra de Lingua e Literatura Galega da Universidade de Santiago: datos para a historia dunha reivindicación” em Brea, M. / Fernández Rei, F., *Homenaxe ó profesor Constantino García*, t. II: 33-53 (Santiago de Compostela: Universidade de Santiago de Compostela).
- Carvalho Calero, R. (1984): “Homenage a Cotarelo”, *Revista Monográfica de Cultura*, 1: 11-12 (A Coruña: A. C. “O Facho”).
- Carballo Calero, R. (1981) “O idioma no Seminario de Estudos Galegos”, *Problemas da Língua Galega* (Lisboa: Sá da Costa).
- Casares C. (1998): “O desastre de 1898 e os escritores galegos”, *Grial*, 137: 5-23.
- Catalán D. (1998): “Permanencia de motivos y apertura de significados: *Muerte del príncipe don Juan*”, *Arte poética del romancero oral. Parte 2ª Memoria, invención, artificio* (Madrid: Siglo XXI).
- Cid J. A. (1992): “*La Colección de los viejos romances asturianos* (1885) de Juan Menéndez Pidal y su formación”, *Asociación Brasileña de Estudios Hispánicos (Abeh)*, 2, 43-73.

- Cotarelo Valledor, A. (1902): *Fray Diego de Deza* (Madrid: Imprenta de José Perales y Martínez).
- [Cotarelo Valledor, A.] (1919): “Peregrinos de un ideal”, *Utreya*, 1: 2.
- Cotarelo Valledor, A. (1919): “Romancero popular gallego”, *Utreya*, 3: 33-37.
- Cotarelo Valledor, A. (1920): “Romance popular gallego. Albucla”, *Utreya*, 18: 277-279.
- Cotarelo Valledor, A. (1921): *Límite septentrional de la conquista sarracena en España* (Santiago de Compostela: El Eco Franciscano).
- Cotarelo Valledor, A. (1923): *Sinxebra* (Santiago: El Eco de Santiago).
- Cotarelo Valledor, A. (1924): *Lubicán. Conto dramático de lobos e de amores, en tres cadros en verso* (Santiago: El Eco de Santiago).
- Cotarelo Valledor, A. (1925): “Un romance galego”, *Nós*, 21: 2-3.
- Cotarelo Valledor, A. (1928): “O pranto da Virxen”, *El Compostelano*.
- Durán, J. A. (1977): *Crónicas, 2. Entre el anarquismo y el librepensamiento* (Madrid: Akal).
- Fernán Vello, M. A./Pillado Mayor, F. (1986): *Conversas en Compostela con Carvalho Calero*, (Santiago de Compostela: Sotelo Blanco).
- Filgueira Valverde, X. (1988): *O arquivo do Museo de Pontevedra* (Pontevedra: Museo de Pontevedra).
- Forneiro Pérez, J. L. (1999): “José Pérez Ballesteros e o romanceliro tradicional galego”, em Álvarez, R./ Vilavedra, D., *Cinguidos por unha arela común. Homenaxe a Xesús Alonso Montero*, t. II: 629-642 (Santiago de Compostela: Dpto. de Filoloxía Galega da Universidade de Santiago).
- Forneiro, J. L. (2002): “Fora as vossas torpes maos do romanceliro galego”, *A Nosa Terra* (5-9-02), 1.046: 28.
- Hervella Courel, A. (1909): *Romances populares gallegos recogidos de la tradición oral* (inédito).
- Juaristi, J. (1994): “El guetto vacío”, em VV. AA., *Auto de Terminación*: 115-136 (Madrid: El País/Aguilar).
- Karady, V. (2000): *Los judíos en la modernidad europea. Experiencia de la violencia y utopía* (Madrid: Siglo XXI).
- Marfany, J.L. (1995): *La cultura del catalanisme* (Barcelona: Empúries).

- Mariño Ferro, X. R. / Bernárdez, C. L. (2002^a): *Romanceiro en lingua galega* (Vigo: Xerais).
- Mariño Ferro, X. R. / Bernárdez, C. L. (2002^b): “Resposta a unha refutación”, *A Nosa Terra* (26-9-02), 1.049: 28.
- Menéndez Pelayo, M. (1900): *Antología de poetas líricos castellanos, romances populares recogidos de la tradición oral*, t. X (Madrid: Librería Hernando).
- Pardo Bazán, E. (1984) [reimp. da ed. de 1888]: “Idioma o dialecto”, *De mi tierra* (Vigo: Xerais).
- Pérez Ballesteros, J. (1884): em VV. AA., *Folklore Gallego. Miscelánea. Biblioteca de las Tradiciones Populares*, vol. IV: 136-137 (Madrid: Fernando Fe).
- Postic, F. (1997): “Le beau ou le vrai ou la difficile naissance en Bretagne et en France d’une science nouvelle: la littérature orale (1866-1868)”, *Estudos de Literatura Oral*, 3, 97-123.
- Rodríguez González, E. (2001): Fernández Valdehorras, C., *Breviario Enciclopédico “don Eladio”. Lendas, historias e tradicións populares de Galicia* (s. l.: La Voz de Galicia).
- Varela Orol C. (1984): “Aproximación á bibliografía de Armando Cotarelo Valledor”, em VV. AA.: *A presenza de Armando Cotarelo en Galicia*: 133-148 (Santiago de Compostela: Dirección Xeral de Cultura da Xunta de Galicia).
- Ventura, J. (2000): *O nacionalismo kármico de Vicente Risco* (Santiago de Compostela: Laiovento).
- Villares, R. (1984): *Historia de Galicia* (Vigo: Galaxia).

Tecnoloxías de la lingua i les seves aplicacións

M. Antònia Martí Antonín

Montserrat Civit Torruella

Mariona Taulé Delor

CLiC-Centre de Llenguatge i Computació

{mtaule,amarti}@ub.edu civit@clic.fil.ub.es

Resumo:

A investigación en Lingüística Computacional e Procesamento da Linguaxe Natural deu lugar estes últimos anos ás denominadas Tecnoloxías da Linguaxe, cuxo obxectivo principal é o desenvolvemento de sistemas informáticos capaces de recoñecer, comprender e xeraren linguaxe humana en todas as súas formas. Con esta finalidade, desenvolveuse unha serie de aplicacións, como a Tradución Automática, a Extracción e Recuperación da Información, a Clasificación de Documentos etc., que procesan a información para facilitaren o acceso, organización e transmisión do coñecemento que xera a chamada Sociedade da Información en que vivimos.

Como noutras disciplinas científicas, na área da Lingüística Computacional e do Procesamento da Linguaxe Natural pasouse dunha etapa inicial centrada na investigación básica de carácter experimental a outra en que se interaxe máis coa sociedade e, por tanto, máis interesada na creación de produtos e aplicacións que resolvan problemas reais. Isto significa desenvolver sistemas e recursos capaces de analizaren a linguaxe sen restricións, isto é, que ofrezan unha ampla cobertura lingüística.

Neste artigo preséntase de xeito introdutorio os recursos (lingüísticos) e as aplicacións máis características que se desenvolven actualmente no marco das Tecnoloxías da Linguaxe. En concreto, salientaremos dos recursos necesarios os analizadores e desambiguadores morfolóxicos e sintácticos, os lexicóns computacionais e os corpus lingüísticos, nomeadamente os etiquetados. Canto ás aplicacións, centrarémonos basicamente na Recuperación e Extracción da Información e na Tradución Automática.

Palabras chave:

Procesamento da Linguaxe Natural, Lingüística Computacional, Recursos Lingüísticos.

Abstract:

In the last years, research on Computational Linguistics and Natural Language Processing has led to Language Technologies, whose main goal is to develop computer systems capable to recognize, understand and generate human languages in all their forms. For this purpose, several applications have been developed, such as Machine

Translation, Information Retrieval and Information Extraction or Document Classification. These applications process the language in order to ease access to knowledge, its organization or its transmission, activities needed by our Information Society.

As in other disciplines, Computational Linguistics and Natural Language Processing have gone from a first period of basic, experimental research to another in which new products and real applications have to be created, in order to solve interaction problems. This means that we need to develop systems and resources capable to deal with unrestricted language, that is, broad-coverage systems and resources. This paper presents an introduction to linguistics resources as well as the main applications being developed nowadays in the Language Technologies framework. More concretely, it emphasizes morphological analyzers, taggers, syntactic parsers, computational lexicons and linguistic annotated corpora. As for applications, stress is laid on Information Retrieval, Information Extraction and Machine Translation.

Key words:

Natural Language Processing. Computational Linguistics. Linguistic Resources.

1. Introducció

El llenguatge ha estat un instrument decisiu en el procés d'humanització que ens ha permès associar-nos, intercanviar informació i transmetre coneixement. Llenguatge i Societat són, per tant, dos conceptes que estan estretament units i no s'entenen ni s'expliquen l'un sense l'altre. En l'evolució de la cultura, l'aparició de l'escriptura constitueix una fita cabdal ja que permet emmagatzemar informació i disposar d'una memòria perdurable enfront del caràcter efímer de la llengua oral. En aquesta línia, la invenció de la impremta constitueix una innovació podríem dir que revolucionària perquè permet la difusió a gran escala del coneixement humà, que influeix decisivament en el procés d'alfabetització i, conseqüentment, a la democratització de la cultura.

A finals del segle XX, la codificació digital de la informació constitueix una nova revolució pel que fa a l'emmagatzemament, organització i transmissió de dades textuals. L'ús generalitzat dels ordinadors personals i, sobretot, l'aparició d'Internet, ha propiciat un creixement exponencial de la quantitat d'informació disponible. Aquestes facilitats en les comunicacions han contribuït a la creació d'un nou nivell de realitat, la realitat virtual, que ens permet realitzar una àmplia gamma d'interaccions. En aquesta nova realitat les distàncies desapareixen de manera que la nostra relació amb l'entorn esdevé global.

La situació actual és, però, d'alguna manera paradoxal: d'una banda, disposem de gran quantitat d'informació susceptible de ser consultada, adquirida i difosa i,

d'una altra, la quantitat d'informació és tan gran i de naturalesa tan variable que es fa molt difícil accedir-hi de manera adient. El resultat ha estat que ha canviat substancialment el concepte mateix del que es considera informació i com tractar-la: no està millor informat qui té més dades, sinó qui disposa dels mitjans necessaris per obtenir específicament aquelles dades que necessita o cerca, aquelles dades que són més apropiades per als seus interessos. És més, el que ara es vol obtenir és coneixement, és a dir, dades organitzades, estructurades i relacionades.

El concepte de Societat de la Informació, és a dir, una societat que es recolza en la producció, organització i transmissió digital de la informació, ha d'incorporar al bé comú que és el llenguatge el desenvolupament de tècniques que facilitin la transmissió del saber, l'organització, interrelació i contextualització de les dades i les parcel·les de coneixement disperses.

L'entorn comunicatiu que deriva de l'aplicació de les noves tecnologies de la llengua planteja reptes fins ara inexistents. Cal dissenyar sistemes informàtics que permetin l'accés a la informació de manera eficaç, ràpida i senzilla. És amb aquesta finalitat que, entre d'altres, les aplicacions de tecnologia lingüística tenen com a objectiu:

- la classificació de documents de manera que es faciliti la seva posterior recuperació;
- el resum automàtic per tal de simplificar-ne la consulta;
- la traducció multilingüe per poder accedir a la informació de documents en diferents llengües;
- el disseny d'interfícies que permetin l'accés a les dades organitzades i classificades;
- l'accés a la informació de manera precisa, com és el cas dels sistemes pregunta-resposta;
- etc.

És desitjable també que s'integrin en un mateix entorn la veu, la imatge i la llengua escrita i que es tingui en compte la varietat de llengües en què es pot trobar la informació, tant per obtenir-la com per poder recuperar els documents que la contenen.

Resoldre aquests problemes exigeix nous plantejaments en el tractament de determinats temes amb llarga tradició en el marc de les tecnologies de la llengua com són la traducció automàtica i la classificació documental, i alhora obre portes al desenvolupament de noves aplicacions com les que acabem d'esmentar. La revolució tecnològica que ha significat l'aparició dels nous sistemes de

comunicació està provocant a més un canvi en profunditat en les professions i sectors productius relacionats amb el llenguatge, com són la traducció, l'edició -en especial de diccionaris i enciclopèdies- i l'ensenyament, que exigeixen una readaptació en profunditat dels professionals d'aquestes àrees.

La possibilitat de dur a terme aquestes noves aplicacions o d'innovar en àrees de coneixement ja existents requereix el desenvolupament de recursos d'enginyeria lingüística que són a la base de qualsevol tractament informàtic del llenguatge humà. Es tracta de sistemes d'anàlisi morfològica, de mòduls de desambiguació morfosintàctica, de bases de coneixement lèxic, de lexicons computacionals, de gramàtiques, d'analitzadors, etc., que són específiques de cada llengua. El desenvolupament d'aquest tipus de recursos requereix de la confluència de diferents agents, públics i privats, atesa la seva complexitat i l'elevat cost de desenvolupament constitueixen el nucli de les anomenades Indústries de la Llengua.

1.1. Les Indústries de la llengua

El Processament del Llenguatge i les Tecnologies de la parla constitueixen la base de l'enginyeria lingüística: l'aplicació dels coneixements sobre la llengua al desenvolupament de sistemes informàtics que puguin reconèixer, comprendre, interpretar i generar llenguatge humà en totes les seves formes.

Encara que el tractament computacional del llenguatge es va iniciar a mitjan dels anys quaranta, no ha estat fins a la darrera dècada del segle XX, als anys noranta, que s'ha iniciat el desenvolupament de sistemes de tractament del llenguatge a gran escala. Aquest canvi d'enfocament ha estat determinat fonamentalment per dos factors bàsics: la difusió de l'ordinador personal, el PC, com a eina habitual per a codificar la informació i l'aparició d'Internet com a mitjà per a transmetre i publicitar la informació. Aquests dos factors han determinat un creixement exponencial de la informació en suport electrònic i, consegüentment, la necessitat de desenvolupar sistemes informàtics per a accedir-hi de manera intel·ligent.

De la mateixa manera que s'ha observat en altres ciències i disciplines, en l'àrea de la Lingüística Computacional i del Processament del Llenguatge s'ha passat d'una primera etapa centrada en la recerca bàsica i en el desenvolupament de sistemes de laboratori de caràcter experimental, a una altra en què l'objectiu és constituir-se en una ciència que interactua amb la societat i que proporciona una base per a la creació de productes de caràcter aplicat.

Recentment s'ha encunyat el terme d'"Indústries de la Llengua" per a designar el complex entramat de productes i serveis que es desenvolupen al voltant del llenguatge: l'edició en suport electrònic, els correctors automàtics, el tractament de

la informació a Internet, el desenvolupament de bancs de dades textuais, etc. precisen de recursos i aplicacions d'Enginyeria Lingüística que facilitin la gestió i l'accés de la informació. En aquests darrers anys, ha aparegut, per tant, una notable demanda d'aquest tipus de serveis i productes lingüístics impensable fa només vint anys, i que ha significat l'aparició de noves professions, necessitats i productes que tenen com a base el tractament informàtic del llenguatge.

Per contra, aquesta nova situació pot agreujar els problemes que tenen plantejats les llengües minoritàries en la mesura que no disposen dels recursos lingüístics necessaris per a participar de ple en la Societat de la Informació. Els criteris de rendibilitat sens dubte decanten l'interès en el desenvolupament de recursos d'enginyeria lingüística per a llengües amb una àmplia difusió i prestigi social i científic. Així, la Societat de la Informació pot actuar com a factor negatiu en el sentit de desequilibrar més la balança a favor d'unes determinades llengües amb una àmplia difusió i amb una presència predominant a Internet, com l'anglès, i a deixar cada cop més en un segon pla les llengües en una situació d'ús més restringida.

Aquest desequilibri es manifesta de maneres molt diverses. D'una banda, s'està agreujant les diferències entre països amb una alta implantació de les noves tecnologies i països en què pràcticament no s'han alterat les maneres de comunicar-se i transmetre la informació. D'una altra, en els països amb una presència rellevant de les noves tecnologies s'està decantant la balança a favor de l'ús de les llengües predominants, per damunt de llengües minoritzades. En les interaccions comunicatives digitals primen els criteris de rendibilitat, cost i impacte de la informació: es tria la llengua que permetrà arribar a més gent ja que els límits d'espai han desaparegut i el missatge pot arribar arreu.

A partir d'ara, entre els factors que es prenen en consideració a l'hora d'avaluar l'estatus d'ús d'una llengua, caldrà també tenir en compte la seva presència en el marc de les tecnologies lingüístiques i l'existència tant de recursos computacionals com d'aplicacions per al seu tractament automàtic.

Si no es vol agreujar la situació en què es troben les llengües minoritzades, s'hauran de prendre mesures que rectifiquin la tendència a la substitució lingüística en determinats contextos d'ús, la qual cosa implica la complicitat de tots els agents implicats que han de facilitar els mitjans econòmics i estratègics per a redreçar aquesta situació.

1.2. Relacions entre la Lingüística i la Lingüística Computacional

La Lingüística Computacional va establir des de la seva aparició una interacció dinàmica amb la Lingüística teòrica que les ha enriquides mútuament. D'una

banda, la Lingüística Computacional va adoptar des d'un bon començament tot l'aparell formal que aportava la Gramàtica Generativa, que incloïa, entre d'altres, el concepte de regles de reescriptura o de producció i el de transformacions; la Lingüística teòrica, d'altra banda, ha obtingut, gràcies a la Lingüística Computacional, la possibilitat d'implementar els models teòrics i desenvolupar un aparell experimental que permet posar-los a prova.

En un terreny més general, el fet de disposar de grans quantitats de corpus en suport electrònic ha permès que els lingüistes tinguin accés a mostres d'ús real de la llengua. Gràcies a les tecnologies de la llengua es poden anotar corpus a gran escala i de manera automàtica a diferents nivells- morfològic, sintàctic i semàntic. A partir dels corpus anotats s'obté informació precisa sobre el comportament de les llengües que permet verificar hipòtesis des de la lingüística i millorar els lexicons i les gramàtiques computacionals. A continuació presentem amb un cert detall el que ha estat la interacció entre la Lingüística i la Lingüística Computacional així com l'impacte de la Lingüística de Corpus en ambdues disciplines.

La publicació el 1957 del llibre *Syntactic Structures* de N. Chomsky i, posteriorment, el 1965, d'*Aspects of one Theory of Syntax* van constituir una aportació de gran transcendència en el desenvolupament de la Lingüística, que va repercutir significativament en altres disciplines com la Filosofia, la Psicologia i la Lingüística Computacional. Chomsky demostrava la insuficiència dels automàtics d'estats finits i de les gramàtiques independents del context com a models generadors de frases d'una llengua natural, i proposava una teoria del llenguatge basada en dos nivells de representació, l'estructura profunda i la superficial, i en dos tipus d'objectes formals: les regles de reescriptura i les transformacions. Per primera vegada una teoria lingüística tenia com a objectiu desenvolupar un aparell formal que permetia descriure i generar les frases de les llengües naturals. Aquesta nova concepció de la gramàtica encaixava perfectament amb els requisits de formalització exigits pels sistemes de processament del llenguatge, que són inherentment formals, ja que els algorismes operen sobre dades discretes de forma que aquells aspectes de les teories que no són formalment expressables no es poden processar.

A partir d'ara i en les diferents formulacions que aniran apareixent, l'objectiu de la teoria lingüística serà caracteritzar totes les possibles llengües naturals i, a la vegada, proporcionar l'aparell formal que permeti la descripció de les llengües en termes de la teoria. Explicar un fet és preveure de quina manera i sota quines condicions es produirà. Així, una gramàtica generativa ha de tenir codificat tot el coneixement lingüístic que permetrà produir totes i només aquelles frases que són gramaticals en una llengua.

Aquests avenços en el terreny lingüístic van tenir una repercussió immediata en el desenvolupament de la Lingüística Computacional a partir de la dècada dels setanta i sobretot als anys vuitanta, en què van començar a aparèixer formalismes i teories sobre el llenguatge amb una clara voluntat de ser processables. Des d'aleshores, l'ús de regles d'estructura de frase, de transformacions, de condicions de bona formació, etc. han estat incorporades en els models computacionals de les llengües naturals. S'havia demostrat en les dues dècades anteriors, en els sistemes de laboratori, que era factible solucionar determinats problemes de processament del llenguatge i ara calia fer un pas endavant, desenvolupar aplicacions reals. Per tal d'assolir aquest objectiu era necessari promoure el desenvolupament de recursos bàsics a gran escala, en especial gramàtiques, lexicons i bases de dades textuals, per al màxim nombre de llengües possible. Com a resposta a aquestes necessitats, al Regne Unit es promou el programa Alvey (Oakley i Owen 1990) i a la Unió Europea s'afavoreix la investigació bàsica i el desenvolupament de recursos lingüístics a gran escala.

Els anys vuitanta es caracteritzen bàsicament per la irrupció de les teories lingüístiques en l'àmbit dels sistemes de processament del llenguatge i per la voluntat de superar les limitacions imposades a partir del concepte de subllenguatge, que limitaven les aplicacions computacionals al tractament de dominis conceptuals i lingüístics restringits (Kitterdige i Lehrberger 1982). El que es pretén a partir d'aquest moment és la creació de sistemes que es caracteritzen per ser capaços de processar qualsevol tipus de text sense cap mena de restricció: les gramàtiques computacionals han de ser de gran abast i tractar una gran varietat de fenòmens lingüístics.

L'exigència de rigor en la formalització del llenguatge és cada cop més gran i els formalismes i teories gramaticals que apareixen al llarg dels anys vuitanta tenen una clara intenció de ser processables computacionalment. Es pot dir que a partir d'aquests moments el tractament computacional és un nou requisit que garanteix que una teoria disposa d'un aparell formal consistent i suficientment expressiu. En aquesta línia, es desenvolupen models lingüístics que tenen com a objectiu captar al màxim les regularitats del llenguatge i així fer generalitzacions lingüístiques el més apropiades possible.

Com a resultat d'aquest enfocament es desenvolupen els formalismes d'unificació o gramàtiques d'unificació, models teòrics o simples anotacions formals, de caràcter marcadament declaratiu, que tenen en comú l'ús d'estructures de trets per a la descripció lingüística i que apliquen l'operació d'unificació per a l'estructuració dels objectes. Entre els primers destaquen la Gramàtica Funcional Lèxica (*Lexical Functional Grammar*, LFG, Kaplan i Bresnan 1982), la Gramàtica d'Estructura Sintagmàtica Generalitzada (*Generalized Phrase Structure Grammar*, GPSG, Gazdar et al. 1985) i la Gramàtica d'Estructura Sintagmàtica Regida pel

Nucli (*Head-Driven Phrase Structure Grammar*, HPSG, Pollard i Sag 1987); entre els segons, cal esmentar el formalisme PATR-II (Shieber 1986), la Gramàtica Funcional d'Unificació (*Functional Unification Grammar*, FUG, Kay 1985), i la Gramàtica d'Adjunció d'Arbres (*Tree Adjoining Grammar*, TAG, Joshi 1984)¹.

La construcció de gramàtiques generals de la llengua planteja el problema de com obtenir informació objectiva i a gran escala, ja que les intuïcions del lingüista no són suficients per a un treball d'aquesta magnitud. En aquest sentit, l'adquisició de coneixement lingüístic es constitueix com una línia de recerca clau per al futur desenvolupament de la Lingüística Computacional per tal que aquesta pugui assolir els objectius que té traçats: la superació de les limitacions imposades pels subllenguatges, la construcció de gramàtiques d'àmplia cobertura i la incorporació de coneixement lingüístic per a la resolució de l'ambigüitat. És en aquest punt que els corpus textuais en suport electrònic adquireixen una importància cabdal ja que esdevenen la principal font d'informació per a inferir coneixement lingüístic, bàsicament per veure de quina manera s'usa una llengua, quin tipus de construccions sintàctiques predominen, quina mena de vocabulari utilitza un determinat domini, en definitiva, per proporcionar la informació necessària per a la construcció de gramàtiques i lexicons computacionals.

1.2.1. La Lingüística de corpus

Un corpus és una mostra d'una llengua que habitualment s'ha construït a partir d'una selecció de textos realitzada segons uns determinats criteris i amb un objectiu concret. Entre els criteris de constitució d'un corpus podem citar: la recopilació de l'obra d'un autor, la representació de la llengua en un moment determinat de la seva evolució, l'exemplificació de determinats fets de la llengua, la selecció de material sobre un domini temàtic específic, etc. Pel que fa als objectius, com acabem d'esmentar en l'apartat anterior, es tracta fonamentalment d'obtenir dades d'ús de les llengües ja sigui tant per a la seva anàlisi lingüística com computacional.

L'existència d'una disciplina específica, la Lingüística de Corpus, que tracta sobre la constitució i processament d'aquest tipus d'objectes, es justifica pel fet de requerir d'una metodologia que difereix en molts sentits de l'emprada per al processament de petites mostres de la llengua. La tipologia de corpus, els criteris per construir-los, les tècniques de processament, la seva finalitat, etc. constitueixen línies de treball clarament diferenciades.

En el seu origen, la Lingüística de Corpus consisteix en l'estudi de les llengües a partir d'exemples d'ús. Els estudis lingüístics basats en corpus no són nous ni tenen

¹ Per a una informació de síntesi sobre formalismes d'unificació es pot consultar A. Moreno Sandoval (2001), M. Taulé Delor i M.A. Martí Antonín (2002) i S. Shieber (1986).

a veure directament amb la Lingüística Computacional: els primers estructuralistes basaven els seus estudis en l'anàlisi de textos. En paraules de Harris (1993): "The approach began ... with a large collection of recorded utterances from some language, a corpus. The corpus was subjected to a clear, stepwise, bottom-up strategy of analysis".

Ara bé, des de fa uns anys, *Corpus* és sinònim de *Corpus en suport electrònic* i *Lingüística de Corpus* ho és de *Lingüística de Corpus en suport electrònic*. L'aparició i l'ús de l'ordinador permet, d'una banda, buscar, recuperar, ordenar i fer càlculs sobre les dades de manera ràpida i eficaç; d'una altra, permet el maneig d'un volum de dades inimaginable fa poc més d'una dècada. La Lingüística de Corpus tal i com s'entén en l'actualitat, està relacionada no tan sols amb l'anàlisi i interpretació de la llengua sinó també amb les tècniques computacionals i la metodologia per a l'anàlisi de textos. La *Lingüística de Corpus en suport electrònic*, segons Leech (1992), focalitza la seva atenció en l'actuació lingüística més que no pas en la competència, en la descripció lingüística, en els models de llenguatge tan qualitatius com quantitativs i en un enfocament de la recerca empirista enfront de l'enfocament racionalista.

Quin és l'interès des d'un punt de vista tant teòric com aplicat, de la Lingüística de Corpus? En primer lloc, cal assenyalar que l'ús de dades externes al parlant, resultat de la producció lingüística, presenta avantatges evidents sobre la introspecció: es tracta de dades directament observables, mentre que la introspecció és subjectiva i opinable. En segon lloc, els corpus són una font d'informació quantitativa important. La freqüència d'aparició dels elements lingüístics no pot recuperar-se per introspecció i si una paraula o una determinada estructura no apareix en un corpus representatiu d'una llengua pot ser una informació significativa, ja que pot qüestionar judicis acceptats per inèrcia. Finalment, els corpus fan possible la verificació objectiva de resultats, especialment quan es tracta de comprovar el bon funcionament d'eines per al processament del llenguatge, però també quan es tracta d'aportar proves per a corroborar les hipòtesis d'una teoria².

Així, el desenvolupament de metodologies basades en l'anàlisi de corpus lingüístics centren l'interès dels estudis entorn del Processament del llenguatge perquè (a) constitueixen una font d'informació valuosa per a la creació de diccionaris i lexicons tant per a l'ús humà com computacional; (b) perquè són una font d'informació fonamental per al desenvolupament de gramàtiques computacionals; (c) perquè degudament anotats a nivell morfosintàctic o semàntic, s'usen en els processos d'aprenentatge automàtic per a construir desambiguadors

² Vegeu per a més informació sobre el tema: Leech (1997), McEnery i Wilson (1996), Ooi (1998) i Civit (2003).

morfològics, sintàctics o semàntics; (d) perquè constitueixen un banc de proves valuós per a evaluar la qualitat dels sistemes d'anàlisi del llenguatge; i, finalment, (e) perquè el propi estudi del llenguatge es pot enriquir gràcies a l'aportació de l'anàlisi de dades textuais.

Sens dubte, una de les aplicacions més exteses i conegudes dels corpus és el seu ús com a font d'informació per a delimitar els sentits dels mots i obtenir exemples, extreure locucions i col·locacions, en definitiva, per a l'elaboració de diccionaris d'ús comú. L'exemple més conegut i pioner va ser el diccionari Collins-Cobuild (Sinclair, 1987), que es va construir íntegrament a partir de la consulta dels exemples existents en un corpus de 6 milions de paraules creat amb aquest fi per la pròpia editorial i la Universitat de Birmingham. Actualment no es concep l'elaboració d'un diccionari que informi sobre l'ús de la llengua sinó és sobre la base d'un corpus representatiu. Els corpus que ara s'utilitzen, però, són de dimensió molt superior, entre els 100 i 200 milions de paraules i normalment s'han etiquetat tant a nivell morfosintàctic com semàntic.

2. Recursos, eines i processos bàsics de Tecnologia Lingüística

Els darrers anys, les línies que han marcat l'evolució de la Lingüística Computacional han estat el desenvolupament, l'estandardització i la compatibilització de recursos lingüístics amb l'objectiu de dotar a les llengües de la infraestructura necessària per a participar en la Societat de la Informació.

Les aplicacions computacionals que tracten informació textual han de fer front a una sèrie de problemes que són característics i inherents a les llengües: la riquesa morfològica i semàntica i l'ambigüitat. Els analitzadors morfològics i sintàctics han de ser capaços de construir representacions formals del contingut dels textos en termes de la seva estructura: han d'assignar la categoria correcta a cada paraula, identificar els constituents i les funcions de les oracions i construir-ne una representació. Les bases de coneixement lexicosemàntic han de permetre associar el sentit correcte a cadascuna de les paraules d'un text i així fer possible la interpretació semàntica de les oracions. Tant en un cas com en l'altre són necessaris programes de desambiguació que seleccionin la categoria morfosintàctica correcta, l'estructura sintàctica adequada i el sentit apropiat de cada ocurrència en el seu context.

Des d'un punt de vista informàtic, les tècniques que es fan servir, i que són a la base dels programes de processament del llenguatge, es basen fonamentalment en dos tipus de coneixement: o bé es recolzen en coneixement lingüístic, normalment expressat en forma d'estructures de dades declaratives i regles, o bé es basen en l'estadística a partir de mostres molt extenses de la llengua, és a dir,

de corpus en brut o bé etiquetats. Els sistemes basats en l'estadística construeixen el seu model de llenguatge inferint-lo a partir dels corpus; en els sistemes basats en el coneixement, en canvi, el model de llenguatge ha estat explicitat prèviament pels lingüistes en forma de gramàtiques, lexicons, bases de coneixement, etc.

En aquest apartat ens centrarem en la descripció dels recursos d'enginyeria lingüística més representatius i que s'usen en els sistemes de processament basats en el coneixement. Alguns, com ara els analitzadors morfològics són a la base de gairebé totes les aplicacions, d'altres, com les xarxes semàntiques tenen actualment un àmbit aplicatiu molt més restringit i s'usen fonamentalment en programes de recerca.

2.1. L'anàlisi morfològica

El component morfològic dels sistemes de processament del llenguatge actua en les primeres fases de l'anàlisi i du a terme una sèrie de tasques que són fonamentals per als processos posteriors d'anàlisi sintàctica i d'interpretació semàntica. És per això, que li dedicarem una especial atenció.

Els analitzadors morfològics reben com a input el text enregistrat en suport magnètic i el resultat que generen és el text segmentat en unitats lèxiques i categoritzat morfològicament. Per tal d'obtenir aquest resultat, l'analitzador morfològic ha de realitzar una sèrie de tasques com són la segmentació del text en unitats lèxiques; la seva lematització, és a dir, l'associació a cadascuna d'aquestes unitats del seu lema i, finalment, l'assignació a cada unitat d'una o més interpretacions morfològiques especificades mitjançant la categoria. El procés d'anàlisi morfològica permet, doncs, relacionar totes les variants flexives d'un mot amb el seu lema i assignar-los informació referent a la categoria i als trets morfològics.

Per exemple, com a resultat d'un procés d'anàlisi morfològica de la paraula *cases*, podem obtenir la següent informació (Figura 1):

Forma	Lema	Informació morfosintàctica
cases	casa	NCFP (nom comú, femení, plural)
	casar	VM2IPS (verb no auxiliar, 2a. persona, indicatiu, present, singular)

Figura 1.

on la paraula *cases* rebra dues interpretacions possibles, una com a nom i l'altra com a verb.

Pel que fa al disseny d'un analitzador, aquest està determinat per diferents factors, entre els quals cal assenyalar:

- les característiques morfològiques de la llengua;
- els tipus de processos morfològics que ha de tractar: flexió, derivació i composició;
- les unitats que es volen identificar: mots aïllats, lexies, temps compostos, perífrasis, etc.

Un analitzador morfològic adient per processar el gallec o el català no necessàriament ha de ser vàlid per processar altres llengües, com per exemple, l'alemany o el turc, tot i que quan es dissenya un analitzador es procura que sigui el més independent possible de qualsevol llengua concreta.

Els analitzadors morfològics es basen normalment en l'estratègia de declarar, d'una banda, les arrels dels mots i els sufixos i, de l'altra, de definir les regles de combinatòria de les unes amb els altres. Es tracta, en darrer terme, de no haver d'introduir una a una totes les variants flexives dels lemes, sinó d'obtenir-les per combinació dels seus morfemes constitutius. Per a llengües poc flexives, com és el cas de l'anglès, una solució habitual per al tractament morfològic és construir un lexicó de formes amb la informació morfològica associada (Vegeu Figura 2).

Lexicó de formes		Expansió del codi
love	V00P, NCS	V00P = verb (V), present (P), no té persona ni nombre (00)
loves	V3SP	NCS = nom (N), comú (C), singular (S)
loved	V00T	V3SP = verb (V), (3) persona, singular (S), present (P)
table	NCS	V00T = verb (V), passat (T), no té persona ni nombre (00)
tables	NCP	NCP = nom (N), comú (C), plural (P)
green	AQ	AQ = adjectiu (A), qualificatiu (Q)

Figura 2.

En el cas de les llengües romàniques, com el gallec o el català, amb una gran riquesa flexiva, en canvi, és recomanable disposar d'un sistema computacional que permeti disposar d'aquest tipus d'informació sense haver de declarar una a una totes les formes de la llengua i la informació associada corresponent. Són precisament els analitzadors morfològics els encarregats de dur a terme aquesta tasca. Podem distingir dos enfocaments possibles en l'estratègia d'anàlisi:

- sistemes d'anàlisi morfològica que actuen a la vegada que analitzen el text: la combinatòria d'arrels i sufixos es realitza en el mateix moment de l'anàlisi;
- sistemes que prèviament generen totes les formes possibles i analitzen el text mitjançant la comparació de les formes del text amb les formes generades automàticament: quan hi ha identitat entre la forma del text i alguna de les formes generades, s'associa a la forma del text la informació continguda en el formari.

Els analitzadors morfològics es basen en la definició de regles que expressen les combinacions possibles d'arrels i sufixos. La construcció d'un analitzador morfològic implica, per tant, definir les classes flexives d'una llengua: cada classe està constituïda per aquelles paraules que presenten el mateix comportament flexiu. Les paraules han de ser analitzades en funció de l'arrel i del sufix (o sufixos) flexiu(s) que admeten. De manera que cada classe d'arrels té un codi que la identifica i els sufixos, al seu torn, s'agrupen en classes segons la categoria que denoten i la classe d'arrels amb què combinen.

Cal tenir en compte, però, que els analitzadors morfològics analitzen textos escrits i els signes ortogràfics no tenen una correspondència exacta amb els sons de la llengua. Així, per al català tenim que els verbs *saltar*, *pelar*, *tallar*, etc. pertanyen a la classe dels verbs regulars de la primera conjugació, mentre que *caçar*, *pegar*, *pecar*, etc. tot i ser igualment verbs regulars de la primera conjugació, presenten canvis ortogràfics i, per tant, s'han de tractar de manera diferent, altrament podríem obtenir com acceptables formes com *caçes (la forma correcta és *caces*) o *menjem (la forma correcta és *mengem*).

Pel que fa a les tècniques d'anàlisi, els autòmats d'estats finits s'han mostrat especialment adients per a la generació i l'anàlisi automàtica de paraules ja que permeten implementar tant les alternances morfoortogràfiques (*cosir/cuso*) com les combinacions d'arrels i sufixos. En aquest marc, hi ha sistemes que implementen un model morfològic de dos nivells (Koskeniemmi 1983), el lèxic i el superficial, que permet tractar les irregularitats sistemàtiques del tipus *caçar/caces* a partir d'una variant ortogràfica que es considera bàsica, mentre que d'altres (Martí 1988) es basen en un model d'un únic nivell que obliga a declarar en el lèxic d'arrels les dues variants ortogràfiques: *caç-* i *cac-*.

Des d'una perspectiva aplicada, la problemàtica de l'anàlisi morfològica és un problema resolt. Existeixen analitzadors morfològics per a llengües molts diverses amb una qualitat del 100% en els resultats.

2.2. Desambiguació morfològica

Els sistemes d'anàlisi sintàctica reben com a *input* els textos segmentats i categoritzats morfològicament. El resultat de l'anàlisi morfològica sol ser ambigu, és a dir, que una mateixa paraula pot rebre més d'una interpretació ja que l'anàlisi morfològica s'aplica sense tenir en compte el context. L'anàlisi sintàctica es pot complicar de manera crítica si el text que ha de processar és morfològicament ambigu perquè s'incrementarà notablement el nombre d'anàlisis possibles. Així, per tal de procedir a l'anàlisi sintàctica o, simplement, per tal d'obtenir una anàlisi morfològica completa i més acurada, cal aplicar un procés de desambiguació en el qual cada paraula rebi només una única interpretació, aquella que li correspon en funció del context en què es troba.

Els desambiguadors morfosintàctics (en anglès, *taggers*) poden basar-se en el coneixement lingüístic o bé en dades estadístiques. En el primer cas, es té en compte el context immediat del mot, les dues o tres paraules que el precedeixen i el segueixen a més d'un conjunt de regles de desambiguació. Un exemple és el cas de la paraula *que*, que pot tenir dues anàlisis possibles: pronom o conjunció. El desambiguador morfosintàctic podria formular les regles següents per determinar en cada cas quina és la categoria més apropiada:

R1: si la paraula que la precedeix és un nom, aleshores és un pronom.

R2: si la paraula que la precedeix és un verb, aleshores és una conjunció

Els desambiguadors que es basen en criteris estadístics resolen l'ambigüitat en base a la solució més freqüent en un determinat context. Aquests sistemes han de disposar prèviament de corpus lingüístics correctament analitzats i desambiguats a partir dels quals poder 'aprendre' les solucions més freqüents donat un context. S'ha comprovat, per exemple, que la paraula *la* és gairebé sempre article, algunes vegades pronom i molt rarament el substantiu *la* (és a dir, la nota musical). Davant de l'ambigüitat els sistemes basats en el coneixement lingüístic decidiran la categoria en funció de les categories veïnes, mentre que els sistemes basats en l'estadística la decidiran en funció de l'opció més freqüent en el context en què es troba a partir de les ocurrences de la paraula *la* en textos prèviament desambiguats a mà. Hi ha també sistemes mixtos que combinen ambdós tipus de coneixement (Màrquez et al. 2001).

Com hem vist, el component morfològic associa informació sobre la categoria i sobre els trets morfològics a cada una de les formes d'un text i, mitjançant l'aplicació de tècniques de desambiguació, s'associa a cada forma un únic lema i una única descripció morfològica. Un cop analitzat, lematitzat i desambiguat, el text es pot sotmetre a un processament d'anàlisi sintàctica i d'interpretació

semàntica o bé es pot utilitzar com a font d'informació per a l'extracció d'informació morfosintàctica i lèxica. Així, es poden obtenir:

- freqüències de lemes i formes;
- freqüències de categories;
- freqüències de seqüències de categories;
- combinacions d'un lema amb determinades categories;
- frases en què apareix un determinat lema o forma;
- etc.

Ara bé, la informació que abasta un analitzador morfològic és limitada. Cal que els mots tinguin també associada informació sobre el seu comportament sintàctic i el seu tipus semàntic per tal de fer possibles altres tècniques de processament del text. Els recursos computacionals que forneixen aquest altre tipus d'informació entorn del mot són els lexicons computacionals i les xarxes semàntiques.

2.3. Lexicons computacionals

La lexicografia computacional té com a objectiu la creació de lexicons que continguin la informació necessària per als sistemes de processament del llenguatge natural i que a més siguin tractables computacionalment. Ha de donar solucions als problemes que planteja la complexitat intrínseca del lèxic, tant pel volum de dades que ha de tractar com per la seva varietat i complexitat.

Des del punt de vista de l'anàlisi computacional, el lèxic constitueix un dels principals problemes que han de resoldre els sistemes de processament del llenguatge natural, ja que de les aplicacions "pilot" dels anys setanta i vuitanta s'ha passat al disseny de sistemes "reals" que tracten les llengües naturals sense cap mena de restricció i això implica poder tractar qualsevol paraula en qualsevol de les seves formes i accepcions possibles.

En els sistemes de processament del llenguatge el component lèxic és una peça clau perquè, d'una banda, permet associar informació a les formes tractades en el component morfològic i, d'una altra, guia el procés d'anàlisi sintàctica i d'interpretació semàntica.

Les unitats de l'anàlisi lexicogràfica són els sentits, és a dir, cadascun dels diferents significats que pot tenir un mot. El fet més habitual és que els mots de les llengües naturals siguin polisèmics, és a dir, que tinguin més d'un significat. El problema, tant de la lexicografia tradicional com dels lexicons computacionals, és delimitar els diferents sentits d'un mot. Cada cop s'accepta més la idea que la identificació

dels sentits està estretament relacionada amb l'objectiu aplicatiu d'un lexicó: els sentits d'un mot variaran si estem construint un lexicó monolingüe o bilingüe, si l'objectiu és construir un sistema de traducció automàtica, d'extracció d'informació o de resum automàtic de textos. De la mateixa manera, en el terreny de la lexicografia tradicional s'observa que les accepcions d'un mot no són les mateixes en un diccionari monolingüe que en un bilingüe, en un diccionari orientat a l'ensenyament de segones llengües o a la consulta general (Martí 2003).

La unitat sobre la qual s'organitzen aquest tipus de recursos és el lema o bé els sentits d'un mot. De fet, la informació que conté un lema o una accepció en un lexicó computacional es considera vàlida per a totes les seves variants flexives. Els lemes són, per tant, les unitats que permeten relacionar les formes d'un text processat morfològicament amb altres dispositius computacionals.

Existeix un cert acord sobre la informació que ha de contenir un lexicó computacional. S'han esmerçat esforços a definir un model d'entrada lèxica de caràcter general a partir del qual es derivin els diferents lexicons, de manera que l'aplicació del model permeti comparar-los, adaptar-los i reutilitzar-los. Es considera bàsica la informació corresponent a:

- el lema o nom de l'accepció;
- la transcripció fonètica, que serà única per a totes les accepcions d'un lema;
- la categoria gramatical;
- l'estructura argumental, o esquema sintacticosemàntic bàsic que caracteritza una entrada verbal;
- les diàtesis, és a dir, els diferents esquemes sintacticosemàntics alternants que accepta l'entrada. Es tracta d'una informació fonamental per controlar el procés d'anàlisi sintàctica i evitar la sobregeneració d'arbres d'anàlisi.
- La informació semàntica, que en els lexicons computacionals s'expressa mitjançant l'associació de l'entrada a un tipus semàntic prèviament definit en una ontologia. Per exemple, a *noi* se li podria assignar el tipus semàntic [HUMÀ] i a *ordinador* [ARTEFACTE]. Els tipus depenen de l'ontologia de què es parteix. En l'apartat següent presentem EuroWorNet, una ontologia que s'utilitza per al tractament de la semàntica en els sistemes de processament del llenguatge. La informació semàntica inclou també l'explicitació de les relacions com la sinonímia, antonímia i la hiponímia/hiperonímia.
- Les equivalències de traducció amb altres llengües.

Els lexicons computacionals no solen contenir tota la informació que hem presentat. La informació bàsica correspon a la categoria i a la llista de subcategorització. A continuació, es dona la possible representació de l'entrada *anar* en un lexicó computacional:

<anar, [[-N], [+V], [BAR 0], [SUBCAT <SP>]]>

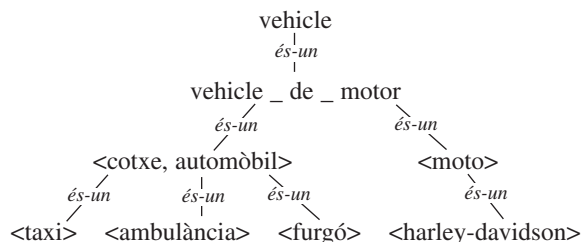
on s'indica que es tracta d'una entrada verbal ([+V]) que subcategoritza un sintagma preposicional ([SUBCAT <SP>]).

Els sistemes de representació més utilitzats per a la representació de la informació lèxica són les bases de dades de caràcter general, les bases de dades lèxiques i les bases de coneixement lèxic, que es poden basar en esquemes (*frames*) o en estructures de trets tipificades (Shieber 1986). Les representacions basades en esquemes o en estructures de trets presenten l'avantatge de poder definir relacions entre les diferents entrades del lexicó a més d'incorporar mecanismes d'herència de propietats que simplifiquen el procés de creació del lèxic i permeten aplicar-hi mecanismes d'inferència de propietats.

Cal esmentar com a representativa d'una nova concepció del lèxic, la teoria del lexicó generatiu de J. Pustejovsky (1995) que defineix un model de lexicó amb un nombre reduït de regles, principis i restriccions que donen compte d'una sèrie de fenòmens lingüístics a un alt nivell d'abstracció. El seu objectiu és constituir les bases d'una teoria del llenguatge de naturalesa semanticolèxica amb capacitat generativa.

2.3.1. Les xarxes semàntiques

Les xarxes lexicosemàntiques codifiquen relacions entre les paraules pel seu significat. Es poden considerar com un tipus de lexicó que se centra en la representació de les relacions lèxiques. Les relacions bàsiques que es tenen en compte són la sinonímia (per exemple, 'cotxe' és un sinònim d' 'automòbil'), la hiponímia ('cotxe' és un 'vehicle de motor') i la meronímia ('roda' és una part de 'cotxe'). Les xarxes lexicosemàntiques relacionen cada paraula, per exemple 'cotxe', amb tots els seus subtipus (o hipònims) -els tipus de cotxes que estan definits a la xarxa- amb el seu tipus semàntic o hiperònim:



Es tracta d'un recurs que s'utilitza per a millorar els sistemes de classificació documental i els sistemes d'extracció i recuperació d'informació: la possibilitat d'accedir als sinònims, als hiperònims o als hipònims d'un mot permet establir relacions entre documents que altrament no es relacionarien, i augmenta a més la capacitat d'interpretació semàntica de les preguntes dels usuaris d'aquests darrers sistemes.

EuroWordNet (Miller i Fellbaum 1991) és una xarxa lexicosemàntica multilingüe que s'està constituint com a un estàndard de la codificació semàntica, en la qual les llengües estan connectades a través d'un índex interlingua de manera que per a cada sentit es poden obtenir les seves traduccions a les altres llengües.³

2.4. Els analitzadors sintàctics i gramàtics formals

Els objectius de l'anàlisi sintàctica són fonamentalment dos: agrupar les paraules en unitats de nivell superior (sintagmes i clàusules) que identifiquin els constituents principals de la oració i etiquetar aquests constituents amb les funcions corresponents. Es tracta de fer explícites les relacions sintàctiques que es troben de manera implícita.

En aquest tipus de procés cal establir una distinció entre l'analitzador o programa informàtic, la tècnica d'anàlisi i la gramàtica. La gramàtica és una especificació formal de les estructures permeses en el llenguatge; la tècnica d'anàlisi és el mètode emprat en el procés d'anàlisi d'una oració per determinar la seva estructura d'acord amb el que s'ha especificat a la gramàtica; finalment, l'analitzador és el programa que, aplicant una determinada tècnica i basant-se en les dades de la gramàtica efectua l'anàlisi sintàctica. Així, en tot programa informàtic que pretengui descriure la llengua hi trobem coneixements de dos tipus: coneixement directament lingüístic i instruccions sobre com fer servir aquest coneixement. Naturalment, el coneixement lingüístic descriu algun aspecte de l'estructura lingüística de la llengua en qüestió. En molts casos, aquest coneixement es formularà en forma de regles d'estructura de constituents com les següents:

1. SN → Det N
2. SN → Det N ADJ
3. SN → Pronom

Pel que fa al nivell de profunditat es distingeixen dos tipus d'anàlisi sintàctica, el parcial o superficial i el total. Les tècniques d'anàlisi total tenen com a objectiu la construcció d'arbres d'anàlisi que representin tota l'estructura sintàctica de l'oració

³ A <http://clic.fil.ub.es/demos/> es poden consultar les xarxes del català, castellà i anglès.

de la forma més detallada possible. Aquests sistemes presenten, però, l'inconvenient de ser molt poc precisos ja que realitzar l'anàlisi sintàctica completa requereix tant informació sintàctica com semàntica i són computacionalment molt lents. Funcionen bé per a dominis específics, ja que tant el lèxic com les estructures que cal tractar són molt inferiors en nombre.

L'anàlisi sintàctica sovint és un procés intermedi en el processament de la informació, la descripció que en resulta sol usar-se posteriorment com a input per a altres processos. És per això que, en la major part dels casos, no es precisa d'una anàlisi en profunditat, fet que ha motivat l'interès a desenvolupar sistemes d'anàlisi parcial o superficial. Aquests sistemes poden identificar constituents i fragments de constituents (en anglès, *chunks*), de manera que es redueix el nombre i tipus de relacions entre aquests elements i no es pretén l'assignació de funcions; conseqüentment, l'ambigüitat es redueix de forma important i l'anàlisi, encara que parcial, conté molt pocs errors. Així en un sintagma nominal com 'la lluita per la vida', s'identificaran dos *chunks*: 'la lluita' i 'per la vida', que corresponen a un sintagma nominal i a un sintagma preposicional. No s'indica la relació entre aquests sintagmes ni es construeix el sintagma nominal complex⁴. Si es vol processar textos no restringits és molt difícil que la gramàtica tingui previstes totes les estructures, elements i combinacions que es poden donar de manera espontània en un text, de manera que es pugui obtenir l'anàlisi completa de cada oració. En aquests casos, l'anàlisi superficial és la solució més recomanable.

L'anàlisi sintàctica pot consistir també en una anàlisi més profunda que permeti extreure i explicitar els constituents i les seves relacions (canvi d'ordre) de manera completa. En aquests casos cal disposar d'una gramàtica de la llengua d'àmplia cobertura i d'un lèxic computacional amb la informació necessària per desambiguar les estructures, identificar els arguments, les funcions, aplicar restriccions selectives sobre els arguments, etc. tal i com s'ha indicat en l'apartat 2.3. Cal disposar, a més, d'un sistema de restriccions que limiti la capacitat d'anàlisi a les estructures correctes. Actualment un analitzador d'aquestes característiques es pot construir per a dominis temàtics restringits en els quals es redueix de manera important l'ambigüitat lèxica i sintàctica i el nombre d'estructures i de fenòmens lingüístics que s'han de tractar.

Actualment s'observa també un interès creixent en el desenvolupament de bancs d'arbres sintàctics (en anglès, *treebanks*) o corpus analitzats sintàcticament (Civit i Martí, 2004). Encara que la seva elaboració pot ser automàtica o manual, només en aquest darrer cas es garanteix la qualitat del seu contingut. En aquest cas, el volum d'aquests corpus és reduït ja que el seu desenvolupament és molt costós. L'interès

⁴ Per al tema de l'anàlisi superficial, vegeu Civit (2003) capítol 4 i Màrquez et al. (2001).

aplicatiu dels bancs d'arbres sintàctics és que constitueixen una font d'informació valuosa per a observar el comportament sintàctic de les llengües i poden servir com a font d'informació per a la inferència de gramàtiques de caràcter general i d'àmplia cobertura.

L'anàlisi sintàctica és encara un tema no resolt tant pel que fa als programes informàtics com als formalismes gramaticals i els models de llenguatge que s'utilitzen. Les Tecnologies de la Llengua han de fer front a l'anàlisi de textos d'ús real de la llengua i, en aquest marc es fa difícil imaginar un programa d'anàlisi que els pugui tractar de manera completa. De moment, l'anàlisi parcial s'ha revelat com una sortida eficient per a resoldre un bon nombre dels problemes que plantegen les aplicacions d'aquestes tecnologies.

3. Aplicacions basades en Tecnologies de la Llengua

Amb el nom d'aplicacions d'Enginyeria Lingüística es designen els sistemes informàtics complexos que resolen tasques intel·ligents utilitzant coneixement lingüístic. Aquestes aplicacions consten de programes (p.e. analitzadors), dades lingüístiques estructurades (com gramàtiques i lexicons) i una estructura informàtica que integra els diferents components.

Atès que el processament del llenguatge és un problema complex d'enginyeria del software i de gestió del coneixement que posa en joc unitats lingüístiques de naturalesa molt diversa, com ara són els morfemes, les paraules, les frases o els significats, i que requereix, a més, l'aplicació de processos tan diversos com ara l'anàlisi morfològica i sintàctica o la interpretació semàntica, és habitual que les aplicacions estiguin dissenyades de manera modular, on cada component s'encarrega d'un tipus de procés.

Les aplicacions d'Enginyeria Lingüística solen classificar-se en dos grans grups segons tractin textos llargs de caràcter narratiu o bé diàlegs. Les primeres presenten la dificultat d'haver de resoldre l'anàlisi de frases llargues, d'estructura molt diversa i difícilment predictable, però tenen l'avantatge de poder resoldre el problema de l'anàlisi a un nivell no excessivament profund i, normalment, només cal que arribin a un processament parcial del text.

Els sistemes dialogats s'usen en entorns de cerca d'informació i, en aquests casos, és necessària una comprensió "completa" del text, ja que altrament seria impossible accedir a les dades que es desitgen. Les frases solen ser més curtes, però no són més fàcils de tractar ja que els textos són més espontanis i, per tant, menys normatius, i sovintegen les referències anafòriques i les el·lipsis, que dificulten la seva interpretació.

L'aplicació més característica del tractament de la informació textual és la traducció automàtica, però actualment estan adquirint un gran relleu altres aplicacions relacionades amb la gestió de continguts, en especial en l'àmbit d'Internet, com són l'extracció i la recuperació d'informació i la classificació automàtica de documents.

Veurem a continuació algunes de les aplicacions més representatives de l'Enginyeria Lingüística.

3.1. La recuperació d'informació

La *recuperació d'informació* (Gonzalo y Verdejo, 2003), una activitat que fins fa poc es restringia a col·lectius de professionals molt concrets -els documentalistes-, ha passat a ser un dels problemes clau a què ha de fer front la Societat de la Informació. Internet es pot considerar com un gran banc de dades que pot devenir inservible si no se'l dota de sistemes de recuperació d'informació que satisfacin les necessitats dels usuaris potencials. El disseny de sistemes de recuperació d'informació s'ha de realitzar tenint en compte que els usuaris d'Internet ja no són un grup reduït d'especialistes, sinó qualsevol individu en qualsevol part del món. A més, les institucions, organismes i empreses que han de gestionar grans volums d'informació són també potencials usuaris d'aquest tipus d'aplicació.

El problema de la recuperació d'informació se centra en dues qüestions fonamentals: el filtrat de les dades i el multilingüisme. Com a filtrat de dades s'entén la selecció d'aquells documents que interessin l'usuari en base a la seva demanda o a uns determinats paràmetres predefinitos: es tracta de recuperar els documents que interessin i evitar el "soroll", és a dir, la recuperació de documents irrelevants. Pel que fa al multilingüisme, cal disposar de recursos computacionals com els que hem presentat al segon apartat d'aquest article que permetin fer les demandes en qualsevol llengua i obtenir els documents en la llengua que l'usuari seleccioni.

De manera general, la recuperació d'informació té com a objectiu recuperar, a partir d'una col·lecció de documents, aquells que corresponen a la cerca de l'usuari. En els sistemes computacionals, la informació textual sol trobar-se emmagatzemada fonamentalment en dos tipus de formats: en forma de dades organitzades de forma tabular (bases de dades) i en forma de textos en el si de col·leccions de documents, com pot ser Internet.

Les bases de dades tenen associats llenguatges d'interrogació que permeten obtenir exactament allò que es desitja, és a dir, tenen sistemes d'accés estandarditzats. En el cas de les col·leccions de documents, el que l'usuari vol recuperar no sol ser una dada específica o un ítem lèxic, sinó més aviat una llista d'ítems, que en aquest cas

seran documents o parts de documents, que tracten d'un determinat tema en què l'usuari està interessat. Els sistemes de recuperació d'informació tracten d'obtenir per mitjans automàtics aquells textos d'una col·lecció de documents que són rellevants respecte d'una determinada pregunta de l'usuari.

Per poder accedir als documents, els sistemes de recuperació d'informació han d'afrontar la realització de les tasques següents:

- Indexar els documents a partir dels ítems d'informació (paraules, expressions, noms propis, col·locacions) que es consideren rellevants. Aquesta tasca es realitza mitjançant l'associació d'un llistat de paraules discriminadores (paraules de l'índex) als documents⁵.
- Representar les consultes de l'usuari en un llenguatge intern interpretable, com a resultat d'haver analitzat la demanda⁶.
- Determinar el subconjunt de documents, els índexs dels quals tenen una alta similitud amb la pregunta de l'usuari un cop ja ha estat interpretada.
- Avaluar la qualitat de la resposta.

El primer d'aquests processos, la indexació, consisteix a assignar a cada document un o més descriptors que el representen i que figuren en un índex predefinit. Quan l'usuari fa una cerca, es recuperen els documents indexats pel(s) terme(s) que conté. El procés d'indexació es pot fer manualment o de manera automàtica. Quan es tracta d'indexar grans volums de documents se sol recórrer a processos automàtics, ja que els resultats entre una i altra metodologia no són gaire diferents.

En el cas de la indexació automàtica cal determinar: a) de quina manera es confejarà l'índex sobre el que posteriorment es farà la cerca; i b) de quina manera s'associa cada document a un o més elements de l'índex. En la confecció de l'índex cal determinar quins tipus d'unitats s'admetran: paraules aïllades, lexies o col·locacions i si s'exclouran o no les paraules funcionals (pronoms, articles, preposicions, conjuncions...).

Si els elements de l'índex s'han simplificat i només hi figuren les formes canòniques dels mots (lemes), cal aplicar tècniques de processament del llenguatge en l'associació dels documents als termes de l'índex. Per a l'anglès, que té una

⁵ En l'apartat corresponent a les tècniques d'extracció d'informació s'expliquen diferents tècniques d'indexar documents.

⁶ Normalment, la interpretació de la consulta de l'usuari implica el seu processament morfològic sintàctic i la seva interpretació semàntica.

morfologia molt pobre, s'apliquen processos d'obtenció de l'arrel (*stemming*) consistents a eliminar els darrers caràcters dels mots que coincideixen amb un element d'una llista predeterminada de morfemes flexius. De manera ideal, la forma resultant hauria de ser un lema, encara que sovint no és així. A les llengües amb més riquesa flexiva s'apliquen processos de lematització basats en l'anàlisi morfològica del text. Alguns sistemes incorporen cert coneixement semàntic i l'índex s'enriqueix amb sinònims i hipònims que amplien la riquesa lingüística de les consultes.

La consulta de l'usuari es representa amb un llenguatge intern interpretable pel sistema. Les preguntes es poden fer a partir de la combinació de paraules clau amb operadors *booleans* (AND, OR), o bé a partir de frases en llenguatge natural. En aquest cas, solen interpretar-se igualment com a paraules clau combinades amb aquests operadors.

La selecció dels documents es pot realitzar de manera automàtica o interactuant amb l'usuari. La selecció automàtica es realitza recuperant aquells textos indexats per les paraules clau que ha expressat l'usuari en la cerca. En els sistemes interactius, es pot incorporar un procediment iteratiu de refinament de la consulta i, d'aquesta manera, s'adquireix el grau de precisió necessari per als objectius de l'usuari.

Des de l'any 1992 tenen lloc les Text Retrieval Conferences (TREC)⁷, congressos de caràcter competitiu que tenen com a objectiu impulsar les tècniques de recuperació d'informació. La base dels TREC és proveir a tots els participants de dades i mètodes d'avaluació uniformes i d'una col·lecció de textos adient per a posar a prova el sistema que volen avaluar. L'aportació d'aquestes competicions és que incorporen mesures objectives per avaluar la qualitat dels sistemes de recuperació d'informació que es posen a prova.

Les línies en què es treballa actualment per millorar aquest tipus d'aplicació és incorporant coneixement lingüístic en la pregunta, especialment informació semàntica i tècniques de tractament dels elements elidits i les referències anafòriques. Una altra línia consisteix a millorar les tècniques d'extracció d'informació per a la confecció de l'índex, que es tracta en l'apartat següent.

Un pas endavant en els sistemes de recuperació d'informació el constitueixen el sistema pregunta-resposta (Vicedo, 2003). En aquest cas el que es tracta de recuperar no és tant un document, sinó el fragment d'un o més documents que donen la resposta precisa a una pregunta de l'usuari. Si bé els sistemes de recuperació d'informació convencionals utilitzen tècniques bàsicament

⁷ <http://trec.nist.gov/>

estadístiques, els sistemes de pregunta-resposta cada vegada més utilitzen tècniques de processament del llenguatge⁸.

3.2. L'extracció d'informació

L'extracció d'informació té com a objectiu processar textos, documents, pàgines WEB, etc. per identificar-ne la informació rellevant, i així, disposar d'informació fàcilment accessible i utilitzable. Com a resultat del procés d'extracció es disposa, en els casos més simples, d'una llista de termes rellevants, que pot funcionar com a resum del propi document o com a índex per als sistemes de recuperació d'informació. En altres casos, s'obté una estructura d'informació organitzada on es recullen les dades més importants.

Les tècniques d'anàlisi que s'apliquen per extreure informació solen ser l'anàlisi morfològica (lematització) i l'anàlisi superficial del text, amb algun component per al tractament semàntic. Alguns d'aquests sistemes fan servir també informació sobre l'estructura del discurs per tal de distingir les parts d'un text que són rellevants de les que no ho són.

Tradicionalment, els sistemes d'extracció d'informació duen a terme el procés havent definit prèviament de manera esquemàtica el tipus d'informació que es vol obtenir. Ens trobem sempre en el marc d'aplicacions que tracten dominis lingüístics restringits que permeten predeterminar el tipus d'informació que es busca. Així, es defineixen plantilles (en anglès *template*) on es representa esquemàticament la informació que s'ha d'extreure: el procés d'extracció consisteix a omplir correctament els diferents camps de la plantilla amb la informació que s'extreu dels documents. En el procés d'extracció es necessita explicitar les relacions semàntiques existents entre els diferents elements seleccionats. Posteriorment, la informació s'incorpora a bases de dades relacionals per a la seva posterior recuperació. Aquestes plantilles poden constituir també la base per a la confecció de resums o bé l'índex per a la posterior recuperació dels documents.

Els congressos del MUC⁹ (*Message Understanding Conferences*), que es realitzen cada dos anys des del 1987, han donat un impuls decisiu al desenvolupament de les

⁸ Es poden consultar sistemes de pregunta-resposta accessibles a través d'Internet:

Omnibase, <http://www.ai.mit.edu/projects/infolab/globe.html>
 IO search engine, <http://www.ionaut.com:8400/>
 Webclopedia, <http://www.isi.edu/natural-language/projects/webclopedia/>
 AskJeeves, <http://www.ask.com>
 LCC, <http://www.languagecomputer.com/>
 AnswerBus, <http://www.answerbus.com>
 Q-GO, <http://www.q-go.com/es/solutions/>

⁹ <http://www.cs.nyu.edu/cs/faculty/grishman/muc6.html>

tècniques d'extracció d'informació. Aquests congressos estan concebuts en termes d'una competició en què els grups de recerca participants posen a prova el seu sistema. Inicialment reben un corpus d'entrenament i, en un màxim de sis mesos, han d'adaptar el seu sistema d'extracció al domini del corpus. Un cop acabat aquest període d'entrenament, es lliura als participants una nova col·lecció de textos i els han de processar amb el sistema ja adaptat. Els resultats de cada grup es comparen amb les plantilles "correctes" confeigides a mà pels organitzadors. El grup guanyador és aquell que ha obtingut les plantilles més semblants a les que serveixen de test. En pocs anys s'espera tenir disponibles aplicacions d'ús general.

Un dels problemes que s'han posat en evidència gràcies als MUC és que hi ha un component important de subjectivitat a l'hora de determinar el que és rellevant en un text i, per tant, no hi ha un acord sobre el que s'ha d'extreure. La coincidència en els resultats en l'extracció manual oscil·la entre el 60% i el 80% i, per tant, es fa difícil d'avaluar la qualitat dels sistemes d'extracció automàtica. Els organitzadors dels MUC han posat un èmfasi especial a determinar mesures objectives, la cobertura i la precisió, per definir el grau de rendiment dels diferents sistemes. Presentem amb un exemple en què consisteix aquesta aplicació de manera paradigmàtica. Suposem que tenim la notícia següent:

El grup INCE va tenir pèrdues netes en el primer trimestre de 2001 de 49 milions d'euros, enfront d'un benefici net de 38 milions en el mateix període de 2000. Aquestes pèrdues es deuen a la seva divisió de cel·lulosa, afectada pel preu de la pasta de paper (un 17% menys)¹⁰.

D'aquesta notícia se n'haurien d'extreure dues plantilles, que corresponen a les dues accions bàsiques que es descriuen, tot i que en la segona el verb és el-líptic. La darrera frase no correspon a una nova acció, sinó que és una explicació complementària i no hi queda representada (Figura 3).

Resultats-1:

Situació: resultat relatiu
Organització: INCE
Signe: negatiu
Quantitat: 49 milions d'euros
Període: 01-01-2001 a 31-03-2001

Resultats-2:

Situació: resultat relatiu
Organització: INCE
Signe: positiu
Quantitat: 38 milions d'euros
Període: 01-01-2000 a 31-03-2000

Figura 3.

¹⁰ Extret de Gonzalo i Verdejo (2003).

L'extracció d'informació té múltiples aplicacions potencials. La més immediata és l'alimentació de bases de dades relacionals que després poden ser consultades i actualitzades amb nova informació. Una altra de les aplicacions més clares dels sistemes d'extracció d'informació és la millora dels cercadors d'Internet (vegeu l'apartat anterior, *recuperació d'informació*).

Els sistemes d'extracció d'informació actuals procedeixen de la combinació de dues tècniques ben diferents: la indexació documental i els sistemes de comprensió de textos de la Intel·ligència Artificial. Podem distingir-hi dues línies metodològiques. D'una banda hi ha els sistemes d'extracció basats en mètodes empírics (estadístics) i, d'una altra, els sistemes basats en el coneixement.

Els primers es recolzen en models estadístics i en tècniques d'aprenentatge automàtic. Segons els documents que s'han de processar, aquests mètodes poden donar molt bons resultats i, a més, tenen un temps de resolució molt ràpid. L'ús de coneixement superficial per a l'extracció d'informació té l'avantatge de ser ràpid d'implementar i que s'adapta fàcilment a nous dominis. Entre els aspectes negatius cal destacar la manca de precisió en els continguts que es recuperen.

Pel que fa als sistemes basats en el coneixement, aquests requereixen la construcció prèvia de fonts de coneixement lingüístic, com són gramàtiques, diccionaris i bases de coneixement. Aquests sistemes tenen el repte de superar l'anàlisi de dominis restringits i sembla que la solució a aquest problema passa per la incorporació de coneixement lingüístic de caràcter general. Alguns models de representació del coneixement lingüístic com les LCS (*Lexical Conceptual Structures*) de Jakendoff (1988) s'estan utilitzant per a aquest fi.

Hi ha una sèrie d'aplicacions que utilitzen tècniques de processament molt properes a les de l'extracció d'informació, algunes de les quals, com la indexació i l'alimentació de bases de dades ja han estat esmentades. Cal esmentar, a més, la classificació de documents, que consisteix a assignar a cada document, d'una col·lecció, una o més categories (classes) en termes d'una classificació predefinida.

3.3. La traducció automàtica

S'entén per traducció automàtica aquella branca de la Lingüística Computacional que s'ocupa del disseny, la implementació, l'avaluació i l'ús de programes informàtics per traduir textos d'una llengua a una altra de manera automàtica. En aquest sentit, la traducció automàtica és l'àrea d'aplicació paradigmàtica de la Lingüística Computacional. Tot i així, és també una de les aplicacions més difícils de tractar, perquè traduir no és simplement substituir paraules d'una llengua a una altra, sinó que implica tot un procés molt més complex en què s'ha de tenir en compte no només la correcta elecció de les paraules en la llengua de destí sinó també cal sovint fer canvis a nivell sintàctic.

Malgrat que la investigació en aquesta àrea avança de manera molt ràpida, les dificultats inherents a la traducció automàtica són moltes. La tasca de traducció implica la resolució de problemes derivats de la diferència estructural entre les llengües, les dificultats per expressar segons quins conceptes en llengües culturalment distants, els problemes inherents de la terminologia tècnica, etc. A aquests problemes cal afegir, a més, aquells que són propis de la traducció automàtica, que deriven de la manca de models complets de comprensió automàtica de les llengües, de la inexistència de models del procés de traducció i de la falta de documentació sobre el tipus d'accions que normalment du a terme el traductor.

Però sens dubte el problema principal de la traducció automàtica és l'ambigüitat inherent del llenguatge natural, ambigüitat que el programa informàtic ha de resoldre a nivell morfològic, lexicosemàntic i sintàctic. No ens hem de sorprendre, per tant, que es tracti d'una aplicació que necessiti disposar de tots els recursos lingüístics que tinguin a l'abast, és a dir, analitzadors morfològics, lematitzadors, desambiguadors, lexicons computacionals, analitzadors sintàctics, gramàtiques, etc.

Hi ha tres factors a tenir en compte a l'hora de dissenyar un sistema de traducció automàtica:

- Si el sistema serà bilingüe o multilingüe (és a dir, si traduirà més d'una llengua).
- Si el mode d'operació serà directa o intervencionista.
- L'estratègia que seguirà, és a dir si es tractarà, bàsicament, d'un sistema de traducció directa, basat en transferència o en memòries de traducció¹¹.

Els sistemes de traducció directa són dels primers que apareixen al mercat i es basen en diccionaris monolingües i bilingües molt grans. Es tracta de sistemes que disposen d'un coneixement lingüístic molt limitat, reduït sovint a un petit mòdul d'anàlisi morfològica i que en cap cas realitza una anàlisi sintàctica. Són, per tant, sistemes molt simples, normalment força ràpids, perquè es limiten a la consulta dels diccionaris, però amb una qualitat de traducció molt baixa.

Els sistemes de traducció basats en transferència són sistemes molt més complexos, en els quals la traducció s'efectua en tres fases: una d'anàlisi, una de transferència i una de generació. En la fase d'anàlisi es procedeix al processament morfològic i sintàctic amb l'objectiu d'obtenir un arbre que descriu l'estructura sintàctica de

¹¹ Deixem de banda els sistemes interllingua perquè actualment gairebé no s'utilitzen. Vegeu (Alonso 2003) per a més informació.

l'oració. És a partir d'aquest arbre sintàctic, que s'inicia la fase de transferència que consisteix a explicitar per a cada element de l'arbre de la llengua font, l'estructura que li correspon a la llengua destinació i a substituir els ítems lèxics de la llengua origen per la seva traducció en base a un diccionari bilingüe. El resultat final de la fase de transferència és un arbre sintàctic on cada paraula de la llengua d'origen s'ha substituït per una paraula en la llengua de destinació i amb algunes indicacions estructurals que s'hauran de tenir en compte en la següent fase. En la fase de generació es du a terme tota una sèrie de tasques pròpies de la llengua de destinació, com per exemple la generació de les formes flexives adequades, la correcta col·locació de les paraules en funció de les regles d'ordre de constituents de la llengua de destinació, inclusió o eliminació de peces lèxiques, etc.

La qualitat dels sistemes de traducció depèn molt del tipus d'informació que s'analitzi, òbviament si entren en joc aspectes que tenen a veure amb la pragmàtica la qualitat de la traducció pot disminuir radicalment, però si en canvi depèn d'informació morfològica, lèxica, sintàctica i, en part, semàntica, es poden obtenir traduccions de qualitat. Un altre dels factors a tenir en compte és la proximitat de les llengües a traduir, els resultats de traducció són normalment millors entre llengües properes (per exemple, entre el català i el castellà o el galleg i el català) que no pas entre llengües molt diferents (com per exemple el català i el rus). De la mateixa manera, és molt important també considerar el tipus de text a traduir, un text tècnic presenta menys problemes que un text d'opinió.

Els sistemes de traducció automàtica es caracteritzen des de fa temps pel tractament de dominis lingüístics restringits o subllenguatges, per la valoració del paper del traductor –que s'ha incorporat als equips de recerca–, per l'acceptació de diferents nivells o graus d'automatització en el procés de traducció i pel desenvolupament d'entorns d'usuari que faciliten la tasca d'actualització de la informació necessària per al funcionament dels sistemes.

Encara que actualment una part de la investigació es dedica al desenvolupament de sistemes que siguin capaços de gestionar diferents dominis lingüístics, no és possible disposar de sistemes comercials de qualitat que siguin capaços de tractar parells de llengües sense cap mena de restricció i, per tant, amb cobertura universal. En aquesta línia, una de les característiques que incorporen alguns dels sistemes actuals és la transportabilitat, definida com la possibilitat de canviar el domini de l'aplicació o les llengües implicades de forma no massa complexa, de manera que l'usuari pugui definir el nou domini o incorporar les dades lingüístiques necessàries sense la intervenció d'un expert informàtic.

Un altre dels aspectes importants dels sistemes de traducció automàtica d'avui dia és la flexibilitat del seu manteniment. La majoria dels sistemes actuals disposen d'entorns interactius per tal que l'usuari –ja sigui lingüista, traductor o lexicògraf–

pugui modificar o augmentar d'una forma àgil i còmoda les dades lingüístiques (especialment les terminològiques) a fi i efecte d'anar introduint les modificacions que milloraran el funcionament del sistema.

Un cop demostrada la inviabilitat dels sistemes de traducció automàtica 'clàssics' desenvolupats en els anys setanta, que pretenien una traducció totalment automàtica i d'alta qualitat per a qualsevol parell de llengües, els objectius d'aquesta aplicació es van orientar cap al desenvolupament de recursos d'ajuda a la traducció, entre els quals es troben els lèxics i les bases de dades terminològiques multilingües. Actualment¹² són també de gran ajut les memòries de traducció, que consisteixen a emmagatzemar de manera paral·lela els textos que s'han traduït i les seves traduccions, tot i establint les correspondències oració per oració i, fins i tot, per segments inferiors al de l'oració. A la llarga, aquesta estratègia permet disposar d'un gran banc de dades de textos traduïts, de manera que davant la traducció d'un nou text el sistema busca a la base de dades quins fragments ja s'han traduït anteriorment i aplica la traducció de manera automàtica.

Enfront dels primers sistemes de traducció que estaven integrats en un entorn de hardware monousuari i es basaven en estratègies de traducció directa, els sistemes actuals funcionen en entorns PC, admeten treballar en forma multilloc de treball i permeten adaptacions personals del lexicó i de les memòries de traducció.

3.4. Les tecnologies de la parla

Les tecnologies de la parla tenen com a objectiu principal facilitar la interacció entre les persones i els sistemes informàtics de la manera més natural possible. En el que hem presentat fins ara, l'objecte de tractament informàtic han estat textos escrits. Ara bé, la manera natural de comunicar-nos és mitjançant la veu i els sistemes informàtics han de ser capaços de superar les restriccions que imposen el teclat i una pantalla. Les tecnologies necessàries per al tractament de la veu són específiques, de manera que constitueixen una disciplina diferenciada respecte del tractament de textos escrits.

El tractament de la parla implica resoldre dos problemes fonamentals: d'una banda, la producció de veu o síntesi i, d'una altra, el reconeixement de veu. El primer, la síntesi, és tècnicament més senzill, tot i que obtenir veus "naturals" és encara un

¹² A continuació s'indiquen les pàgines web on es poden consultar els diferents sistemes que hi ha de traducció automàtica del català, inclosa la variant valenciana:

Internostrum, <http://www.torsimany.ua.es/>

SALT, <http://www.cult.gva.es/dgoiepl/salt/>

AutomaticTrans, <http://www.automatictrans.es>

Compendium, <http://compendium.es>

problema no resolt. El segon, el reconeixement de la parla, és un tema més complex ja que el sistema informàtic ha de ser capaç de segmentar la cadena fònica de la manera correcta la qual cosa implica tractar un gran nombre de variables a més d'altres factors com poden ser el soroll, errors, etc.

La síntesi de veu consisteix a generar un senyal vocal, és a dir, parla, a partir d'un sistema informàtic que utilitza dades i regles introduïdes prèviament (Llisterri 2003). Els conversors de text en parla, constitueixen un subtipus d'aquests sistemes, i el seu resultat és l'equivalent a una lectura en veu alta d'un text. En l'actualitat es disposa de sistemes que permeten generar missatges verbals en diferents condicions, el problema és que no presenten la naturalitat pròpia de la parla humana. Les línies de recerca actuals s'orienten a millorar la naturalitat dels conversors text-veu incidint especialment en la prosòdia i a millorar la seva flexibilitat tractant que es reflecteixi l'estat emocional que hauria d'acompanyar el text. Es tracta de qüestions no menyspreables ja que, de vegades, constitueixen l'únic mitjà de comunicació de les persones que han perdut la capacitat de comunicar.

El reconeixement de la parla consisteix en la generació automàtica d'un conjunt de símbols discrets (en general, un text escrit) a partir del senyal de parla (Llisterri 2003): es tracta, per tant, de convertir el senyal fonètic en una representació que pugui tractar un sistema informàtic. Els sistemes de reconeixement es caracteritzen per tres paràmetres bàsics: el tipus d'enunciats que poden tractar, el nombre de locutors que accepta un sistema i el vocabulari que es capaç de reconèixer l'aplicació.

Els sistemes de reconeixement de la parla poden tractar tant paraules aïllades com la parla continua., el que s'anomena dictat automàtic. En el mercat ja es troben productes que permeten aquesta funcionalitat, però presenten encara limitacions: solen ser sistemes que necessiten fases prèvies d'aprenentatge (l'usuari ha de proporcionar al sistema mostres de la seva veu per tal que la pugui reconèixer), la seva qualitat depèn de si tracten dominis restringits, i no estan exempts d'un percentatge d'error.

L'àmbit d'aplicació dels sistemes de reconeixement de la parla pot ser tant domèstic com industrial, i resulta especialment útil per a persones amb limitacions de mobilitat ja que els facilita la interacció en un entorn prèviament acondicionat: amb la veu es poden encendre i obrir els llums, pujar i abaixar es persianes, etc. També es poden usar aquests sistemes per a resoldre tasques en què és habitual dictar la informació, com ara els informes mèdics, els documents legals, etc. En determinats contextos, poden substituir el que fins ara s'ha dut a terme amb una gravadora, amb l'avantatge que posteriorment no cal transcriure el text.

Les línies de recerca en reconeixement de veu s'orienten a millorar la qualitat dels resultats, i a inserir-los en el marc d'aplicacions basades en el diàleg, com poden ser les consultes telefòniques a centres d'informació, els serveis d'informació ciutadana o la traducció automàtica per telèfon, on es combina amb d'altres tecnologies, tant de la imatge com del text.

3.5. D'altres aplicacions

Fins aquí hem presentat les aplicacions més característiques de les Tecnologies de la Llengua. Es tracta però d'una àrea de recerca i desenvolupament dinàmica que està generant noves aplicacions a mesura que cal donar resposta a les noves necessitats que planteja la pròpia dinàmica de les Indústries de la llengua. Podem citar, en aquesta línia, la classificació de documents i la producció de resums, camps d'aplicació propers a l'extracció d'informació.

En la classificació de documents l'objectiu és assignar a cada un dels documents que s'han de tractar una o més categories d'entre les que proposa una ontologia o una taxonomia. Els classificadors automàtics es basen habitualment en la presència de determinades paraules representatives o característiques que permeten associar el document a la classe.

Els sistemes de resum automàtic solen consistir sovint en una llista d'expressions simples o d'expressions multiparaula que el sistema ha considerat rellevants com a representatives del seu contingut. El problema és determinar quines són aquestes frases o expressions. Els primers sistemes feien la selecció basant-se en criteris de freqüència. Posteriorment s'han tingut en compte aspectes referents a l'estructura del document o del propi text per donar més rellevància a determinats fragments. Aquests sistemes recolzen fonamentalment en mètodes estadístics, tot i que actualment s'incorpora cada cop més coneixement lingüístic per tractar determinades formes d'ambigüitat, resoldre les anàfores i la correferència.

Els beneficis d'aquestes aplicacions són evidents: la classificació automàtica de documents, associada a un sistema de flux massiu d'informació com pot ser Internet, permet seleccionar els documents interessants segons un determinat perfil, i així tenir-los disponibles per a la seva posterior recuperació. La confecció de resums ens estalvia la tediositat d'haver de llegir molts textos quan tractem de trobar, dins d'un domini, els que tracten d'un tema específic. Com fàcilment es pot endevinar, ambdues tècniques, associades amb l'extracció d'informació, tenen un gran futur en la gestió de continguts a l'àmbit de les grans xarxes d'informació.

Bibliografía

- Alonso, J.A. (2003): “La traducción automática”, dins *Tecnologías del lenguaje*: 94-129 (Barcelona: Editorial UOC).
- Chomsky, N. (1957): *Syntactic Structures* (The Hague: Mouton).
- Chomsky, N. (1965): *Aspects of one Theory of Syntax* (Cambridge, Mass.: The MIT Press).
- Civit, M. (2003): *Criterios de etiquetación y desambiguación morfosintáctica de corpus en español. Monografías*, 3, de la SEPLN (Alicante: Sociedad Española para el Procesamiento del Lenguaje Natural).
- Civit, M. / Martí, M. A (2004): ‘Building Cast3LB: a Spanish Treebank’, a *Treebanks and Linguistic Theories*, Kluwer Academic Publishers (en preparación).
- Gazdar, G. / Klein, E. / Pullum, G. / Sag, I. (1985): *Generalized Phrase Structure Grammar* (Oxford: Basil Blackwell).
- Gonzalo, J. / Verdejo, F. (2003) “La extracción y recuperación de información” dins *Tecnologías del lenguaje*: 157-192 (Barcelona: Editorial UOC).
- Harris, Z.-H. (1993): *The Linguistics Wars* (Oxford: University Press Oxford).
- Jakendoff, R. (1988): *Semantics and Cognition* (Massachusetts: MIT Press).
- Joshi, A (1984): ‘How Much Context-Sensitivity is Required to Provide Reasonable Structural Descriptions: Tree Adjoining Grammars’, a Dowty, D. / Karttunen, L. / Zwicky, A. (eds.), *Natural Language Processing: Psycholinguistic, Computational and Theoretical Properties*: 190-205 (New York: Cambridge University Press).
- Kaplan / Bresnan (1982): ‘Lexical-Functional Grammar: A Formal System for grammatical Representation’, a Bresnan (ed.), *The mental Representation of Grammatical Relations*: 173-281 (Cambridge, Mass.: MIT Press).
- Kay (1985): ‘Functional Unification Grammar: A Formalism for machine Translation’, a *Proceedings of COLING 84*: 75-78 (California: Menlo Park).
- Kittredge, R. / Lehrberger, J. (1982): *Sublanguage. Studies of language in restricted Semantic Domains* (New York: Walter de Gruyete).
- Koskeniemmi, K. (1983): *Two-Level Morphology: A General Computational Model for Word-Form Recognition and Production*. PhD Thesis, University of Helsinki, Department of General Linguistics.
- Leech, G. (1997b): ‘Introducing Corpus Annotation, a Garside / Leech / McEnery (eds.), *Corpus Annotation. Linguistic Information from Computer Text Corpora*: 1-18 (London: Logman).
- Llisterri, J. (2003) ‘Las tecnologías del habla’, dins *Tecnologías del lenguaje*: 249-281 (Barcelona: Editorial UOC).

- Martí, M.A. (1988): *Processament informàtic del Llenguatge Natural: un sistema d'anàlisi morfològica per ordinador*. Tesi Doctoral, Dept. Filologia Romànica, Barcelona.
- Martí, M.A. (2003): "Consideraciones sobre la polisemia", dins Martí, M.A / Fernández, A. / Vázquez, G. (eds.): *Lexicografía computacional y semántica*: 61-103 (Barcelona: Edicions Universitat de Barcelona).
- Martí, M. A. (coord.) (2003): *Tecnologías del lenguaje* (Barcelona: Editorial UOC).
- Martí, A. / Castellón, I. (2001): *Lingüística Computacional* (Barcelona: Edicions Universitat de Barcelona).
- Màrquez, LI / Padró, LI / i Rodríguez, H. (2001): 'Mètodes robustos en l'anàlisi del llenguatge. El processament de text no restringit, *Lingüística Computacional*: 1-68 (Barcelona: Edicions de la Universitat Oberta de Catalunya).
- McEnery, G. / Wilson, A. (1996): *Corpus Linguistics* (Edinburgh: Edinburgh University).
- Miller, G. A. / Fellbaum, Ch. (1991): "Semantic networks of English", *Cognition*, 41: 197 -229.
- Moreno Sandoval, A. (2001): *Gramáticas de unificación y rasgos*. Col. Lingüística y Conocimiento. 32 (Madrid: A. Machado Libros, S.A.).
- Oakley / Owen (1990): *Alvey, Britain's Strategic Computing Initiative* (Cambridge, Mass: MIT Press).
- Ooi, V. B. Y. (1998): *Computer Corpus Lexicography* (Edinburgh: Edinburgh University press).
- Pollard, K. / Sag, I. (1987): *Information-Based Syntax and Semantics*. Volume 1: *Fundamentals*. CSLI, Lecture Notes, 13. (California: Stanford).
- Pollard, K. / Sag, I. (1993): *Head-Driven Phrase Structure Grammar*. CSLI (Chicago / London: The University of Chicago Press).
- Pustejovsky, J. (1995): *The generative lexicon* (Cambridge MA: The MIT Press).
- Shieber, S. (1986): *An introduction to unification-based approaches to grammar*. CSLI, Lecture Notes, 4 (California: Stanford).
- Sinclair, (1987): *Looking Up* (London: Collins ELT).
- Taulé, M. / Martí, M. A. (2001): 'Formalismes gramaticals', *Lingüística Computacional*: 1-99 (Barcelona: Edicions de la Universitat Oberta de Catalunya).
- Vicedo, J. L. (2003): *Recuperación de información de alta precisión: los sistemas de búsqueda de respuestas*. *Monografías*, 2, de la SEPLN (Alicante: Sociedad Española para el Procesamiento del Lenguaje Natural).

Os topónimos *Acea de Ama* e *O Xurés*

Gonzalo Navaza
Universidade de Vigo

Resumo:

Este artigo ocúpase da historia e o significado de dous nomes de lugar galegos: *Acea de Ama* e *O Xurés*. O primeiro, denominación dunha localidade dos arredores da Coruña, é un topónimo composto do arabismo *aceá* ‘muíño’, a preposición *de* e o hidrónimo prerromano *AMA*, que é o antigo nome da Ría do Burgo. O arabismo *aceá*, que perdeu a consoante nasal etimolóxica, proporciona interesante información acerca dalgúns cambios fonéticos na historia do galego e a súa cronoloxía.

Tanto o nome galego (*O Xurés*) coma o portugués (*Gerês*) dos coñecidos montes da raia seca proceden doutro hidrónimo prelatino que identificamos cos *Originis*, *Oregines* e *Ocerensis* rexistrados no *Itinerarium Antonini* e no mapa chamado do Ravenate, formas latinizadas ou deturpadas dun **OGERENSE* ou **UGERENSE* orixinario.

Palabras chave:

Toponimia galega. Hidronimia galega. Hidronimia prerromana. Arabismos.

Abstract:

This paper studies the history and the meaning of two Galician placenames: Acea de Ama e O Xurés. The first one, the name of a little village near A Coruña, is a composed toponym containing the arabism acea ‘mill’, the preposition de and a pre-latin hydronym AMA, which is the ancient name of the Ría do Burgo. The arabism acea, which lost the etymological nasal consonant, provides interesting information about some phonetical changes in the history of Galician language and its cronology.

*Both the Galician and Portuguese names of the well known mountains in the Portuguese frontier called O Xurés (O Gerês) come from another pre-latin hydronym which we identify with the words Originis, Oregines and Ocerensis (written in the Itinerarium Antonini and in the map called Ravenate), latinisated or mistaken forms of an original *OGERENSE.*

Key words:

Galician toponymics. Galician hydronymics. Pre-latin hydronymics. Arabisms.

1. Acea de Ama

1.1. Tradición escrita e interpretacións etimolóxicas do topónimo

Durante boa parte do século XX, *Haciadama* foi a forma escrita oficial do nome dunha entidade de poboación próxima á cidade da Coruña, pertencente na actualidade ás parroquias do Burgo e de Rutis no concello de Culleredo. Madoz e outras fontes do XIX rexistraban habitualmente este topónimo como *Aciadama*, mais no século XX (polo menos desde o Nomenclátor de 1910) fixouse como oficial a forma con *h-* inicial, que se mantivo vixente durante máis de setenta anos nos escritos públicos e na rotulación viaria. Tal grafía obedecía seguramente a unha etimoloxía popular que vía no topónimo un composto da preposición castelá *hacia* (< lat. FACIE AD ‘cara a’) e o substantivo común feminino *dama*. Nas etimoloxías populares dos topónimos, coma as que recolleu Xosé María Álvarez Blázquez (Baión, 1976) e outras moitas que seguen vivas na tradición oral á espera dunha recolleita sistemática, non é insólita a pretensión de explicar os nomes galegos a partir da lingua castelá. Neste caso a interpretación do topónimo como “Hacia+dama” puido vir reforzada pola proximidade do lugar chamado *Ombre*, que tamén viña escribíndose habitualmente con *h-* desde tempo atrás¹ e que moveu a fantasía das xentes a entender a parella de *Haciadama* e *Hombre* como dous topónimos dialogantes ou replicantes.

Durante o longo período da súa vixencia como forma oficial, a grafía *Haciadama* suscitou a curiosidade e os comentarios de xornalistas e eruditos. Eugenio Carré Aldao, na *Geografía general del Reino de Galicia* dirixida por Carreras Candi, desacreditaba o *h-* do topónimo identificando fundadamente o seu primeiro elemento co substantivo común galego *aceá* ‘muíño’, e suxería a posibilidade de que o segundo elemento fose o adxectivo correspondente ao latín MALA:

Acea da Má (Aceña de la Mala), que en parte es del Burgo y en parte de Rutis y que en la rídícula toponimia oficial se convirtió en Haciadama [...] Máa es final de muchos nombres toponímicos. No sabemos qué significará en ellos. En gallego equivale a mala (Carré 1980: 757 e nota 585).

¹ Xa se rexistra coa grafía *Hombre* no documento de 1435 que citamos máis adiante. No século XIX Acea de Ama e Ombre pertencían á mesma parroquia de Almeiras, segundo consta en Madoz (1845). O topónimo *Ombre* repítese varias veces na xeografía galega. Carré Aldao combateu a opinión que identificaba co castelán *hombre* o nome da parroquia de Ombre, en Pontedeume, e defendeu unha etimoloxía vasco-céltica (Carré 1980: 757), rexeitada por Moralejo Lasso, quen tampouco aceptaba a hipótese de Luís Monteagudo segundo a cal *Ombre* sería deturpación castelanizante dun hidrónimo prelatino *Ome* (Moralejo 1977: 325 e 83). Isidoro Millán explica o *Ombre* eumés como composto céltico en *-bre* (<-BRIG) a partir do radical hidronímico prelatino *UM- (Millán 1987: 141-142). Máis recentemente, Edelmiro Bascuas defende para algúns *Ombre* galegos e portugueses este mesmo radical, mais cun segundo elemento diferente, a partir dun étimo *UMERI con *b* epentético, como en *Tambre* < TAMARE. Non sería esta, así e todo, a orixe dos *Ombre* de Pontedeume e do Pino, para os cales a documentación medieval obriga a considerar un étimo ANOBRE, formado co mesmo segundo elemento defendido por Millán xunto co radical hidronímico prelatino ANA- (Bascuas 2002: 235-236).

Na década de 1970, Abelardo Moralejo Laso retomaba a explicación de Carré Aldao e aceptábaa nos seus dous termos, confirmando o primeiro cun documento de Sobrado de finais do século XII onde se fai referencia a unhas aceas construídas nese lugar ou nas súas inmediacións por homes do mosteiro:

Un caso curioso de topónimo desfigurado y contestado por razón de su *h* inicial y otras razones es *Haciadama*, nombre de una aldea entre El Burgo y Rutis en Culleredo (Coruña), que fue anteriormente *Acea da Má* (o *Máa*) “Aceña de la Mala”, según la *Geografía*, Coruña, I, pág. 757, y cuyo origen atestigua una cita documental de A. del Castillo, *Notas a la hist. antig. de La Coruña*, 1948, pág. 48, «illas *zenias* quas fecistis iuxta Burgum de Faro in rivulum qui discurrit iuxta Villam Bonam (hoy Vilaboa)» (a. 1186). *Acea* es común en gallego y aparece como topónimo una docena de veces, alguna en plural, en Coruña, Lugo y Orense, aunque no es normal su evolución desde el árabe *as-saniya* por la pérdida de *n* ante *yod*. Tal vez por eso no figura en la *Toponimia árabe de España*, de Asín Palacios, figurando en cambio *Aceña* o *Aceñas* en Lugo, Orense y Pontevedra, aparte algunas otras provincias. En el topónimo, actualmente disfrazado con la *h* y la grafía como un solo vocablo, ha habido además reducción de una sílaba por reducción del hiato a dip-tongo, y de la doble acentuación a única, que si no ha variado debe ser aguda. Pero habrá que atenerse a la pronunciación de la gente del lugar (Moralejo 1977: 325-326, nota 29).

A atención concedida por Moralejo a este topónimo, difundida nas aulas, as conferencias ou os xornais polos lectores de *Toponimia gallega y leonesa*, avivou a finais desa década as protestas contra a forma deturpada, con frecuencia acompañadas da defensa da proposta etimolóxica de Carré Aldao, *Acea da Má*, agora aparentemente corfirmada polo catedrático compostelán. Na páxina de cultura de *La Voz de Galicia* do 6 de maio de 1979 salientábanse as formas “Haciadama” e “Niño de la Guía” (esta por *Niñodaguia*) como casos “de barbarie maior cometidos contra a nosa cultura en xeral e coa nosa toponimia en particular”², e propugnábase para a primeira, mediante unha explícita fotografía do rótulo de *Haciadama* riscado e corrixido, a grafía *Acea da Dama*. Esta nova hipótese, foneticamente verosímil (por haplloxía das sílabas *da-da*) pero igualmente carente de fundamento documental, seguía a recomendación de Moralejo de aterse á pronuncia local, xa que a realidade oral do topónimo supoñía un grave atranco para a defensa da forma oxítóna. Non o viron así outros estudiosos, como González Garcés, que, se ben concordaba coa consideración de *Haciadama* como “aberrante e insultante”, non admitía a proposta de *Acea da Dama* e seguía mantendo a hipótese de Carré Aldao, supostamente confirmada por Moralejo, *Acea da Má* (González Garcés 1987: 149).

² Non indaguei a identidade dos dous responsables da páxina, cualificados de “lusistas distinguidos” por González Garcés (1987: 149), de quen tomo a referencia.

Na década de 1980 o concello de Culleredo acometeu a normalización da súa toponimia e no nome de *Haciadama* procedeu a emendar as dúas chatas que segundo Moralejo disfrazaban o topónimo (o *h* e a grafía como un só vocábulo), optando pola proposta de Carré. Deste xeito a modificación para *Acea da Ma* ou *Acea da Má*, que non é só gráfica senón que contén un desllocamento da tonicidade, acabou incorporándose ao nomenclátor oficial da provincia da Coruña na edición de 1986 e continuou vixente durante tres lustros. Esa era a forma que ata hai pouco empregaban os medios de comunicación para se referiren ao lugar, e facíanlo con frecuencia debido ás instalacións que ten alí o Deportivo da Coruña.

Que semellante alteración do nome oficial dunha localidade non tivese ningún tipo de contestación pode entenderse polas condicións poboacionais desa localidade, que soporta a presión demográfica expansiva da cidade da Coruña, de xeito que a poboación autóctona é proporcionalmente pouco significativa. E non só nas últimas décadas. Hai máis de século e medio, o dicionario de Pascual Madoz, entre a sucinta información que daba sobre o concello daquela chamado de Alvedro, salientaba a existencia de segundas residencias de coruñeses. Na entrada correspondente á parroquia de Almeiras, tras enumerar as entidades de poboación que a integran (“Aciadama, Almeiras, Alvedro, cap. del ayunt., Hombre, Marisqueira, Telba y Vigo-Vidin”) engade que tiña “80 casas, varias de recreo, en las que habitan por temporada algunos vecinos de la Coruña”. Haberá que atribuír a estas condicións o insólito triunfo da forma inventada polos etimoloxistas, que, de prosperar, engadiría un exemplo toponímico aos casos de palabras radicalmente alteradas por causa dun erro da etimoloxía erudita³.

No ano 2000, a Xunta de Galicia, seguindo o ditame da Comisión de Toponimia apoiado en fontes documentais, estableceu como oficial a grafía de *Acea de Ama*, e con esta grafía figura no *Nomenclátor de Galicia. Toponimia oficial das provincias, concellos, parroquias e lugares*, publicado pola Xunta no ano 2003.

1.2. Aceas e muíños

A voz *aceas*, que ten como étimo remoto o árabe *a-sāniya* ‘roda hidráulica’, constitúe un exemplo de arabismo con presenza garantida na nosa lingua desde o período protorromance. Recóllela a maior parte dos dicionarios do galego moderno,

³ Dadas as circunstancias que vive o noso país no momento de redactar estas liñas, vén ao caso lembrar como exemplo de alteración provocada por un erro etimolóxico erudito o da familia léxica derivada do latín *PRESTIGIUM*. Cando na Francia do XVIII o común *prestige* foi perdendo o seu significado etimolóxico e panrománico de ‘ilusión, xogo de mans, engano’ para adquirir o de ‘creto ou reputación positiva’, o seu parente léxico *prestigiateur* foi interpretado erradamente polos etimoloxistas como un composto de *preste* ‘rápido’ e *digitu* ‘dedo’, e dese erro naceu *prestidigitateur*, logo adoptado e adaptado por outras linguas.

onde se rexistran tres acepcións diferentes: unha máis xenérica, como sinónimo de muíño (ou propiamente de muíño hidráulico); outra máis específica, que distingue as aceas dos muíños hidráulicos en razón da roda vertical das primeiras, e unha terceira restrinxida, que dá o nome de *aceá* a esa roda vertical. A máis antiga fonte lexicográfica galega que recolle esta voz son as papeletas de Sobreira, onde figura coa segunda destas acepcións:

ACEA. S.f. Especie de molino a cuya rueda mueve la corriente del agua estando perpendicular = Aceña. *Lat.* Moletrina aquaria. *Arab.* Cinia. (Santamarina 2001 [Sobreira 1787]: s.v.)⁴.

Asígnalle ese mesmo significado o primeiro dicionario da Real Academia Galega (1913-28), que explica as diferenzas:

La aceña y el molino ordinario se diferencian, en que aquella, es movida por una gran rueda colocada verticalmente, sobre cuyas paletas obra el agua al caer, y el molino se mueve por medio del rodicio (rodeznó).

Seguen este modelo Carré Alvarellos (1928-31) e Eladio Rodríguez (1958-61). Carré Alvarellos engade ademais a terceira das acepcións mencionadas:

ACEA s. f. Aceña, molino harinero que mueve el agua por medio de una gran rueda colocada verticalmente. Esta misma rueda (Santamarina 2001 [Carré 1974]: s.v.).

Outros dicionarios, probablemente máis fieis ao significado que adoita presentar o vocábulo *aceá* na fala, non distinguen nidiamente entre esta e o muíño hidráulico común. O de Francisco Xavier Rodríguez (1863), ademais de deixar constancia do uso toponímico do termo, facía estribar a oposición entre a *aceá* e o muíño nunha vaga diferenza de “aguante”:

Acea. 1. Aceña, especie de molino de más aguante que los comunes. 2. Hay varios lugares con este nombre, supongo por haber o haber habido junto á ellos, aceas (Santamarina 2001 [Rodríguez 1863]: s.v.).

Outras fontes establecen novas oposicións, ás veces diverxentes segundo as áreas xeográficas. Así, en *Frapas III*, Elixio Rivas recolle:

Acea. s. f. Molino que tiene cubo; el pequeño, sin cubo, es albeiro. Cabreiros de Xermade, Lu. En Vilaseco de Cea, Our.: *aceá*, muíño de rodicio horizon-

⁴ As referencias dos dicionarios galegos, desde Sobreira e Francisco Xavier Rodríguez ata o *Frapas III* de Elixio Rivas, tómoas do *Diccionario de Diccionarios* (Santamarina 2001). Do de Leandro Carré reproduce a 4ª edición, de 1972 (A Coruña: Moret).

tal, tamén dito *bruia*. En Barxamaior do Cebreiro, Lu.: *aceña* é un muín onde pineiran; os que non pineiran son *muíños*. En Noceda do Courel: *aceña*, molino. Hai eilí unha nesa corrente que eínda moe (Santamarina 2001 [Rivas 2001]: s.v.).

En portugués, malia a pretensión de eruditos e lexicógrafos de precisaren as diferenzas, *azenha* tamén pasou a adquirir o significado xenérico, equivalente a *moinho*. No século XVIII explicaba Viterbo que a *azenha*

difere do moinho: este tem rodízio e anda com água do rio, e aquela tem roda pela parte de fora, e anda com água de ribeiro que, caíndo d'alto na roda, lhe dá o impulso. Ficou-nos este nome dos Árabes, que chamam *assanha* ao moinho d'água, que serve para trigo. E nós hoje, ampliando a sua significação, chamamos azenhas não só as que moem o pão mas também as que pisam a azeitona. [...]

[...] nos princípios da nossa monarquia (quando os *moinhos de azas* ou de vento, apenas inventados, se não praticavam ainda em Portugal), *azenha* e moinho eram sinónimos, prescindindo do diverso artefacto com que uns e outros eram construídos (Viterbo 1993, s.v. *azena* [sic]).

Se o desenvolvemento dos muíños eólicos restituíu para *azenha* o seu significado de “moinho que anda com água”, o mesmo volveu suceder en datas máis recentes, coa aplicación de novas enerxías como a eléctrica. Deste xeito restableceuse a diferenza entre *moinho* e *azenha*, especificando para esta a condición hidráulica:

azenha s.f. moinho de rodízio movido a água (Do ár. *as-sániâ*, ‘nora; roda de irrigação’) (Porto Editora 2002: s.v.).

Azenha. s.f. (998 cf. JM³) moinho movido a água. ETIM ár. *as-s,niya* ‘roda de irrigação’; no port.ant., var. *acenha*; f.hist. 998 *azenia*, 1365 *zenha*, sXIV *azenha*, 1440 *azenhar* (Houaiss 2001: s.v.).

En castelán, o dicionario da Academia asígnalle ao substantivo *aceña* unha primeira acepción co significado de “Molino harinero de agua situado dentro del cauce de un río” e unha segunda como sinónimo de *azud*, isto é, roda hidráulica (DRAE, s.v. *aceña*). O catalán *sínia*, *sénia* e o aragonés *cenia* non teñen o significado de ‘muíño’, senón exclusivamente o de ‘noria’, mentres que en éuscara se rexistra *azenia* como nome do rodicio ou roda inferior dun muíño hidráulico (Corominas DCECH, s.v. *aceña*).

Nas fontes consultadas non rexistramos para o portugués *azenha* o significado restrinxido de roda hidráulica, mais é de crer que tamén existiu esa acepción etimolóxica, dada a presenza que ten noutras áreas peninsulares. En galego, como se viu, só a rexistra o dicionario de Leandro Carré Alvarelos. Podería pensarse que Carré

o toma dalgunha fonte castelá. Na documentación galega antiga a voz correspondente ao actual *acea* adoita designar sempre un muíño, polo que cabe supor que, malia o exposto por lexicógrafos e eruditos, puido dárselle ese nome a calquera muíño hidráulico, independentemente, como en portugués, “do diverso artefacto com que uns e outros eram construídos”. Só encontramos en galego unha única referencia a unha *acea do muíño*, que talvez aluda á roda. É un documento de Allariz do século XIV:

pela mea **da azea do moyño** et pelas meas das casas en qu eu morava [...] as quaes Azea et casas et Leyras fforon da dita Mna Eans mia moller (Díez Sanjurjo 1906: 95).

1.3. Cronoloxía da palabra e da cousa

Aínda que os muíños movidos por enerxía hidráulica xa eran coñecidos en Roma, a xeneralización do seu uso na maior parte da Península constitúe unha innovación medieval difundida desde al-Andalus, e co seu espallamento difundíuse tamén a voz arábiga que os designaba. No territorio galego, arredor do ano 1000 o cereal mófase habitualmente co muíño “de una sola muela accionado por enerxía humana o animal”, segundo Carlos Baliñas Pérez, quen afirma que ese debe ser o significado que ten na maior parte dos casos o *molinum* ou *molendinum* dos documentos galegos anteriores a esa data (Baliñas 1998: 121). Con posterioridade, *muíño* acabaría por ser denominación xenérica de todo tipo de muíños, incluídos os hidráulicos, pois a medida que estes se foron facendo comúns tamén foi facéndose innecesaria a precisión semántica que levaba a denominalos cunha palabra distinta da usada para designar os primitivos muíños manuais.

González Garcés, a propósito do citado documento de 1186 que fai referencia ás aceas (“illas zenias”) construídas polo mosteiro de Sobrado nas Mariñas dos Freires “iuxta Burgum de Faro in rivulum qui discurrit iuxta Villam Bonam”, cualifica os muíños hidráulicos de “maquinaria adelantada para su tiempo” (González Garcés 1987: 149 e 463-464). Por máis que estes enxeños non coñecesen aínda a difusión que logo habían de ter, non podemos admitir a afirmación do erudito coruñés. A propia existencia do nome galego *acea* é sinal de que non eran unha novidade nesa época de finais do século XII, cando a caída do -N- intervocálico en galego estaba máis ca consumada, senón que debían levar tempo entre nós, aínda que non fose dun xeito xeneralizado. Segundo Carlos Baliñas:

desde finales del siglo IX se documentan entre las pertenencias de la *villa* auténticos molinos fluviales, fuente de extensos rendimientos económicos para los grandes propietarios señoriales, que son casi los únicos que pueden realizar las inversiones necesarias para su construcción (Baliñas 1998: 121).

Cerca dun cuarto de milenio antes da referencia ás aceas do mosteiro de Sobrado na coruñesa ría do Burgo, entre as propiedades dunha doazón ao mosteiro de Celanova figura un muíño localizado “in suburbio Colimbrae” no ano 950:

medietatem de villa Anzana integra, ipsa medietas cum suo molino, qui annis singulis exsolvet molinus ipse frumenti modios XL, hordei similiter XL, simulque milio modios XL qui sub uno sunt modii numero centum XX (Andrade 1995: I, 27).

Un muíño que moe anualmente unhas 30 toneladas de cereal⁵ ten que ser un muíño hidráulico. Se o mosteiro de Celanova no ano 950 era propietario de polo menos un deses artefactos, é doado supor que contribuiría á construción de muíños semellantes (ben de roda vertical, ben de rodicio interno) en terras da súa propiedade situadas moito máis ao norte, isto é, no territorio galego actual.

Noutro documento celanovense datado dúas décadas antes (ano 932), entre unhas propiedades adquiridas polo mosteiro cerca do Sil, en terra de Quiroga, lemos:

aquas cursiles vel incurbiles secundum in alias cartas vestras delimitant de Sancto Petro usque in Ranarios, et usque in Ganderas, cum omnes suos exitus, sua roboreda et sua salta de terminos de Miron usque in carrale antiqua, sive et suas pescarias in ipso flumine, et suos sessos de molinos in ipso arroyo qui discurrit de ipsa villa (Andrade 1995: I, 231).

Eses muíños situados nun curso fluvial han de ser por forza muíños hidráulicos, talvez construídos a imitación dos modelos máis meridionais.

Coa chegada do obxecto debeu chegar tamén a súa denominación, de xeito que a forma galega *acea* sería un dos primeiros arabismos asumidos polo léxico do noroeste. A data da súa incorporación pódese relacionar con acontecementos históricos como a toma de Coimbra (c. 878) por Mendo Guterres (Hermenegildus Goterici), onde teñen orixe algunhas das propiedades de Celanova en terras do Mondego, e cos movementos demográficos desa época en que se produce a inflexión definitiva da transición entre o latín e o romance, estudada por Amable Veiga (Veiga Arias 1983), unha época durante a cal se asentan no noroeste peninsular importantes grupos de cristiáns procedentes do sur islamizado, chegados das rexións de Toledo (top. Toldaos), de Coimbra (top. Cumbraos, Cumbráns), de Viseu (top. Bixueses < VISEUENSES?)... Poderíamos datar a chegada ao noso territorio das *aceas* e da palabra que as designa entre finais do século VIII e o primeiro tercio do século X.

⁵ O dicionario de Eladio Rodríguez atribúe ao *moio*, no seu uso como medida para cereais, o valor de doce ferrados (“en algunhas comarcas gallegas”). Segundo Leandro Carré, o ferrado de millo tiña 20,87 litros.

1.4. *Aceas* e *aceñas*

A variante *aceña* non figura nos dicionarios galegos, non sendo nos máis recentes, seguramente pola súa coincidencia co castelán, que levou a interpretala como forma allea á nosa lingua. A presenza de *aceña* na toponimia galega tamén é significativamente menor, e fronte á ducia de exemplos que Moralejo sinalaba para *aceá* (localizados nos concellos de Castro de Rei –onde se rexistran tres–, Allariz, Baños de Molgas, Betanzos, Cedeira, Culleredo, Monfero, Muros, Ortigueira e Sobrado, aos cales debemos engadir *Veiga da Acea* en Castro de Rei e *A Graña da Acea* en Monfero, ademais dun topónimo urbano *Acea de Olga*, en Lugo⁶), só se rexistran tres topónimos maiores coa forma *aceña*, un en singular, *A Aceña* (no concello lugués de Ribas de Sil) e os outros dous co plural *As Aceñas* (en Ribadeo e Vilanova de Arousa).

Aceña é un resultado regular en galego do árabe *a-sāniya* ‘muíño movido por enerxía hidráulica’, coincidente coas formas castelá e portuguesa e documentado en galego desde a Idade Media, concorrendo con *aceá*. En documentos anteriores ao século XIV o castelán rexistra por escrito, ademais da forma común *aceña*, *açēña*, unha variante con sonora na sibilante, *azeña*, e esa mesma vacilación existiu no galego medieval, como veremos, e persiste no portugués contemporáneo, onde, malia a defensa que algúns lexicógrafos como Cândido de Figueiredo fixeron da forma con xorda (Corominas DCECH: s.v. *aceña*), finalmente prevaleceu na lingua *padrão* a variante con sonora, *azenha*.

Na toponimia portuguesa é un tema moi abundante. No *Dicionário corográfico* de Frazão contamos máis de 70 localidades chamadas *Azenha* ou *Azenhas*, espalladas por todo o territorio, e un só exemplo con xorda, *Assenha*, no concello de Leiria (Frazão 1981: 65)⁷.

A vacilación da sibilante vémolada tamén en textos galegos de comezos do século XIII: “acenia”, “azenas e moynos”, “las azennas e hos moyños” (Lorenzo 1977: 878, s.v. *moyno*). En xeral predomina a grafía correspondente á sibilante sonora, o mesmo ca no latín *zenias* do documento do século XII citado por Moralejo.

⁶ Para este conviven as denominacións *Aceña de Olga* e *Acea de Olga*; se ben a preferida como forma oficial no ámbito municipal é a segunda, non posúo información suficiente que confirme que se trata dunha forma histórica ou se é, como sospeito, unha alteración de *Aceña de Olga*, erradamente interpretado como castelanismo o seu primeiro elemento. O segundo corresponde ao substantivo común galego *olga* (paralelo do cast. *huelga*), celtismo que o latín tardío rexistra como *OLCA* ‘terra de labor’.

⁷ Dada a tendencia uniformadora da toponimia oficial portuguesa, que nos topónimos transparentes tende a acomodar as grafías ás formas do *padrão*, independentemente das realizacións locais, cabe a posibilidade de que algunhas das *Azenhas* actuais correspondan a formas tradicionais con sibilante xorda (por exemplo, localizamos na Internet unha herdade chamada “da Assenha” en Santiago de Cacém). Ademais das formas mencionadas, existe un topónimo *Assanhas*, no concello de Celorico da Beira, mais non dispomos de información que permita confirmar a súa pertenza a este tema léxico.

Se *aceña* constitúe un resultado normal en galego do árabe *a-sāniya*, a forma *acea*, como advertía Moralejo, ofrece a aparente dificultade da perda da consoante nasal ante iode. Na lingua medieval, aínda cando o común nese contexto é a palatalización da consoante nasal, non son infrecuentes os casos en que se produce a súa perda, polo que *acea* sería un exemplo próximo aos medievais *comuion*, *testemoio* ~ *testemuio*, *soterraio*, etc., que convivían cos resultados con nasal palatal (*comuñón*, *testemuño*, *soterraño*...). Ferreiro suxire a posibilidade de que esa evolución diverxente se deba, nos casos de perda da nasal, a unha influencia erudita (Ferreiro 1995: 173). Se é así, habería que supor que a tendencia cultista operou só nun primeiro momento, en época temperá do período protorromance, e impediu nalgunhas voces a palatalización de n+iode, mais nun segundo momento, cando se produce a perda xeral do -N-, eses vellos cultismos pasarían a comportarse coma as voces patrimoniais que perden a nasal intervocálica.

Para a cronoloxía do fenómeno da caída de -N- en grupo con iode interesa saber que no campo da onomástica persoal encontramos unha evolución fonética paralela á de *acea* no nome feminino medieval *Ougea* < EUGENIA, precedente dos actuais *Uxia* e variantes. Se ben este nome de orixe grega xa tiña presenza nos repertorios cristiáns da Alta Idade Media, na onomástica peninsular divulgouse especialmente a partir do culto a Santa Ouxea (Sancta Eugenia Cordubense), mártir a mans de Abderramán III, coma o noso San Paio, e igual ca el canonizada e promovida como obxecto de devoción desde o mesmo século X.

Con todo, fronte á posibilidade de que a perda da nasal no grupo con iode sexa un fenómeno exclusivo de termos incorporados á lingua nese período (dos séculos VIII ao X) en que se divulgou no noroeste a forma *acea*, debemos considerar a existencia dalgúns hidrónimos galegos de orixe prelatina formados cun sufixo -ENIA (así os documentados en latín medieval como *Elenia* ou *Armenia*) que presentan hoxe formas sen consoante nasal palatal: *Ella*, *Armea*. Edelmiro Bascuas, que se ocupou destes hidrónimos (2002: 45 e 293), cre que a dupla evolución de -N- ante iode obedece a unha vacilación temperá na silabación: a secuencia vocálica ou ben é monosilábica (e hai, xa que logo, un iode que pode palatalizar a consoante nasal precedente) ou ben bisilábica (e entón hai un -N- intervocálico).

Ao contrario do que acontece con *aceña* (que se ben é resultado regular en galego tamén podería ser unha achega posterior, tomada dos romances veciños), a forma *acea*, coa síncope da consoante nasal, proba que o termo entrou no noroeste peninsular no período en que esa síncope era activa. Como é sabido, a caída de -L- e -N- orixinouse no territorio da antiga Gallaecia e desde aquí “propagou-se, com a Reconquista, ao Centro e Sul de Portugal, onde as referidas consoantes se conservavan” (Maia 1986: 568).

O portugués ten soamente as dúas variantes mencionadas (con xorda ou con sonora na sibilante, mais sempre con nasal palatal) e non coñece unha forma equivalente

do galego *acea* no léxico común, o que non significa que non a coñecese nalguna época da súa historia, polo menos en áreas miñotas e trasmontanas. David Lopes non inclúe ningún exemplo sen consoante nasal entre os descendentes toponímicos portugueses do étimo árabe que el translitera como *açania*, e rexistra unicamente casos de *azenha* (Lopes 1968: 38). Non podemos pronunciamos sobre o nome da freguesía chamada *Azias*, no concello de Ponte da Barca (isto é, en territorio da Gallaecia) sen contar con documentación antiga que confirme o seu parentesco co galego *acea*. Menos probabilidade lle vemos a *Azia*⁸ (lugar do concello de Algezur, no Algarve), que parece corresponder ao rexistrado indistintamente como *Asaia* ou *Asia* no ano 1220, segundo vemos en Cortesão, quen recolle ademais outro topónimo *Assaiola* (ano 906, localización xeográfica non identificada) que semella un derivado diminutivo (Cortesão 1912: 34).

1.5. O hidrónimo *Ama*

O segundo elemento do noso topónimo non é o “da Má ou Múa” suxerido por Carré Aldao e comentado por Moralejo. A pronuncia local, que o profesor compostelán non chegou a verificar, nunca foi oxítona. Contén a preposición *de*, sen artigo, seguida dun elemento que a documentación antiga rexistra coa forma *Ama* e que interpretamos como un hidrónimo que daría nome ao curso de auga en que estaba construída a acea. González Garcés, comentando o citado documento de 1186 onde se fai mención das aceas de Sobrado, localízalas “en el riachuelo que va desde Vilaboa a la Ría del Burgo”. Cabería pensar que *Ama* era o nome primitivo dese “riachuelo”, pero polos testemuños antigos debemos considerar que *Ama* era o nome do que hoxe coñecemos como a Ría do Burgo, onde conflúen o río Mero e outros cursos de auga menores. O lugar do Burgo, que acabou impondo o seu nome á ría e condenou ao esquecemento a denominación máis antiga, aparece ocasionalmente na Idade Media como *Burgo de Ama*. De se conservar esta denominación, o topónimo *Acea de Ama* tamén estaría menos exposto ás deturpacións, pois, tal como explicaba Jud, a *metacedeuse* ou cambio de parentela é enfermidade propia dos elementos que quedan orfos de familia léxica.

No testamento de Afonso Fernández, *jurado da vila da Crunna*, datado en 1435 aparecen repetidamente estes topónimos:

Iten mando meu Casal de Vigo Vedín con suas casas e cortiña e con a viña da agra da Iglesia e con **outra viña de Ama** [...] e con metade do **Bacelar da Zeadama** como se de parte do rego de longo contra o camiño que bay para o **Burgo de Ama**... (González Garcés 1987: 615, doc. nº 63).

⁸ Estes topónimos figuran coas grafías *Asia* e *Asias*, respectivamente, no *Dicionário Corográfico de Frazão*. A grafías oficiais son hoxe *Azia* e *Azias*.

Neste texto, que contén a documentación máis antiga que coñecemos de *Acea de Ama*, o *a-* inicial de *azea* aparece deglutinado, interpretado como artigo, aínda que non puidemos comprobar se figura así no documento orixinal ou é unha segmentación do transcritor.

Ama é un topónimo de orixe prelatina formado sobre o coñecido radical hidronímico indoeuropeo *AM- ‘canle’ (Hoz 1963: 236). Elixio Rivas asocia a este radical topónimos como *Amaía*, *Amio* e *Ames*, e vincula atinadamente con eles o fitónimo *amieiro* / *ameneiro* “árbore amiga das correntes de auga” (Rivas 1994: 114 e 158). A documentación antiga obríganos a excluír desa familia hidronímica prerromana o nome do concello de *Ames*, a pesar de que a súa localización xeográfica leve a relacionalo co nome da comarca chamada *A Amaía* ou *A Maía*: rexístrase nos textos medievais como *Oiames*, *Oyames* e *Oames*. Semella igualmente un topónimo prerromano, pero non contén o mesmo radical *AM-. Nun documento de 1295 vemos xuntos *Ames* e *Amaía*: “friigesia de sam Thome d’Oyames en terra d’Ameya” (FDU: 18).

Edelmiro Bascuas inclúe o segundo elemento do topónimo *Acea de Ama* entre os parentes de *AM-, raíz que tamén estaría presente nos nomes do *amieiro* ou *ameneiro*. Para este autor, *Ama* constitúe a formación máis simple das que conteñen esta raíz hidronímica, e, se ben os derivados mediante sufixo que ofrece son relativamente abundantes (*Amio*, *Amaía* ou *A Maía*, o *Maia* portugués e o *Amaya* burgalés, *Meixe*, *Amoexa*, *Amoeiro*, *Amoedo*, *Molgas*, *Amoroce*, etc. ademais dos fitotopónimos relacionados co *amieiro*), da súa forma máis elemental só recolle o noso *Ama* e dúas atestacións medievais: un *Amo* en Curtis e un diminutivo *Amela* cerca de Sobrado, documentados nos séculos XII e XIII, respectivamente, e non identificados na actualidade (Bascuas 2002: 208-209).

Podemos engadir dous novos exemplos toponímicos co segundo elemento de *Acea de Ama*. O primeiro é un lugar da parroquia de Samieira, no concello pontevedrés de Poio, chamado Riodama (= Río de Ama), que toma o seu nome da corrente que o rega. Sarmiento rexistra este topónimo na súa viaxe de 1745: “Samieira. Río de Ama. Aren. Ragió. Priorato de Armenteira” (Pensado 1975: 58), e no *Catálogo de voces* explica: “El río de Samieira se llama de Ama” (Pensado 1973: 277). Luísa Blanco transcribo erradamente como “O río Dama” (1995: 77).

O segundo exemplo é un pequeno curso de auga chamado *Regueiro de Fontedama* en Meira (Moaña)⁹. Debemos analízalo como *Fonte de Ama*, cun primeiro elemento *fonte* que alude ao lugar onde nace o regueiro.

⁹ Entre as aldeas de Couso e O Pombal (folla 223-I do mapa IGN 1:25.000). Agradezo a localización exacta ao meu alumno Iván Sestay, que me proporcionou as coordenadas do topónimo: 42°18'0"N, 8°43'25"W.

O indoeuropeísta Francisco Villar, na súa descrición do sistema vocálico paleohispánico, inclúe os hidrónimos *amanus*, *amanum* xunto con *umanum* como exemplos do que el chama “terceira oscilación” (Villar 2000: 369-371). Se así fose, a raíz *AM- sería unha variante diverxente doutra raíz hidronímica *UM-, de recoñecida presenza na hidronimia galega (*Uma*, *Umia*, *Mao* < UMANO..., etc.).

1.6. O nexo preposicional ante hidrónimo

O único elemento do composto *Acea de Ama* que ten orixe latina é o nexo preposicional. De admitir a graffía actual do galego o emprego dos apóstrofos, unha forma escrita como *Acea d’Ama* reproduciría máis fielmente a pronuncia local dese nexos.

Nalgunhas fontes moi recentes úsase en lugar da preposición unha contracción desta co artigo feminino, *Acea da Ama* (por exemplo, na páxinas dedicadas ao concello de Culleredo na web da Deputación da Coruña¹⁰). Así mesmo, no documento coruñés de 1435 onde aparecen referencias a este topónimo, atopamos rexistrados, xunto con “Viña de Ama” e “Burgo de Ama”, outros exemplos en que o hidrónimo figura aparentemente con artigo feminino: “Riba a Ama”, “Riba da Ama” (Gonzalez Garcés 1987: 613). Igual que sucede con outros topónimos compostos en que o segundo elemento introducido pola preposición é un hidrónimo, a forma tradicional habitual é sempre sen artigo, aínda que en hidrónimos comezados por vogal haxa tendencia a deglutinala ao interpretala como artigo.

Os documentos antigos rexistran sempre os compostos sen artigo no hidrónimo, como segue a ser habitual en toda a toponimia galega: *Ribas de Sil*, *Parada de Sil*, *Barra de Miño*, *Castrelo de Miño*, *Salvaterra de Miño*, *Xinzo de Limia*, *Ribadelouro*, *Ribadetea*, *Ridadumia*, *Riba de Neira*, *Ribadavia*, *Santa Cristina de Asma*, *Antas de Ulla*, *Freáns de Deva*, *Xunqueira de Ambía*, *Santa María de Sar*, *Pazos de Arenteiro*, *Barxa de Lor*, *Santa Baía de Búbal*, *Couso de Salas*, etc., e o mesmo sucede noutros dominios lingüísticos peninsulares: *Miranda de Ebro*, *Aranda de Duero*, *Paredes de Coura*, *Ponte de Lima*...¹¹.

¹⁰ <http://www.dicoruna.es/municipios/Culleredo/geografia/localizacion/>

¹¹ Son excepción na toponimia portuguesa os nomes de *Valença do Minho* e *Miranda do Douro*, sen dúbida relativamente recentes, que aluden no segundo elemento non só ao río senón tamén á rexión ou territorio administrativo, dun xeito paralelo a outros topónimos como *Celorico da Beira*. O primeiro rexístrámolo na documentación máis antiga unicamente como *Valentia*, *Valencia* e *Valença*, sen segundo elemento; só a partir do século XVIII comeza a aparecer como “do Minho”. Na toponimia galega existen algúns casos semellantes en que o hidrónimo aparece sempre con artigo: *Grañas do Sor*, *Val do Dubra*, etc. Este último é invención moderna que substituíu o nome do concello chamado Buxán (Buján) no século XIX. Na toponimia portuguesa tamén se pode observar na actualidade certa tendencia a incorporar un artigo ao nome do río, como sucede en *Ponte de Lima*, que ocasionalmente vemos escrito como Ponte do Lima.

En *de Ama, de Avia*, etc., a vogal inicial dificulta a identificación do nexa, que pode ser interpretado polo escribán ou o paleógrafo como contracción da preposición co artigo. Fóra destes casos, na documentación antiga os nomes dos ríos por vía de regra non van precedidos de artigo.

Así mesmo, sintagmas comúns na lingua actual como “o Río Miño” ou “o Río Deza” (artigo + xenérico + hidrónimo) non son habituais na documentación medieval, que en contextos equivalentes emprega a preposición entre o xenérico e o hidrónimo:

*et en fondo fere eno **Rio de Auya*** (ano 1393, Ribadavia, 72)
*una pesquera q nos tenemos eno **rio de miño*** (ano 1451, Almeida, 84)
*Porto de Felgueyra, que he eno **rio de Miño*** (ano 1493, Devanceiros, 208)
*e vay por esa coredoira no **rio de Deua*** (ano 1343, Ramirás, 494)

Foi o uso dos sintagmas sen preposición (fórmula nova ou innovadora, talvez castelanizante ou polo menos favorecida pola coincidencia co castelán) o que posibilitou, por elisión do xenérico *rio*, denominacións como “o Miño”, “o Deza”, “o Ulla”, propias da lingua actual pero insólitas na toponimia, que é máis conservadora. Así e todo, a crenza de que as variantes con artigo son máis galegas levou en tempos recentes a empregar formas como *Parada do Sil, Salvaterra do Miño, Xinzo da Limia, Colexiata do Sar* ou *Terra do Deza* en lugar das tradicionais (e oficiais) *de Sil, de Miño, de Limia, de Sar, de Deza*. Vemos a continuación algúns exemplos da lingua medieval. Nótese que os topónimos actuais *Ribas de Miño* e *Ribas de Sil* tiñan antigamente en singular o primeiro elemento:

*autem decimas in Burgo **de Avia*** (ano 1197, Oseira I: 106)
*Migel Viviançe notario **de Deça*** (ano 1280, Oseira 2: 1094)
*en Bayona & no val **de Miñor*** (ano 1280, Maia: 187)
*a Junqueyra **de Limia*** (ano 1281, Maia: 133)
*moesteyro de Santesteyro de Riba **de Sil*** (ano 1288, Devanceiros: 43)
*notario jurado en Riva **de Mino*** (ano 1288, Oseira II: 1146)
*feleglesia de sancta Ougea **de Asma*** (ano 1290, Oseira II: 1157)
*do concello de Bayona **de Miñor*** (ano 1296, Maia: 206)
*en terra **de Deça*** (ano 1305, Oseira II, 1257)
*da friiGRESIA de santa Maria **de Sar*** (ano 1349, FDU: 209)
*Santa Christina de Vilarino de Riba **de Mino*** (ano 1356, Rocas: 202)
*que he em terra **de Deça*** (ano 1365, Devanceiros: 95)
*su o sino de Santa Cristina de Riba **de Miño*** (ano 1401, Espadanedo: 175)
*San Byeyto de Riba **de Miño*** (ano 1493, Devanceiros: 209)
*coto d’Abeleeda del Val **de Salas*** (ano 1516, FDU: 562)

2. O nome da Serra do Xurés / Gerês

2.1. Unha proposta arabista errada

No primeiro dos artigos titulados “O elemento árabe en galego” aparecidos nesta mesma revista, Xavier Frías Conde expón algunhas consideracións xerais sobre os arabismos na nosa lingua (Frías 2000: 157-171). Trátase dun campo insuficientemente atendido, como sinala o autor. Desde os escritos de Sarmiento, para quen a ausencia do estrato árabe era case axiomática e constituía o *mellor* trazo caracterizador do galego fronte ás linguas veciñas (Navaza 2003), os estudos sobre o tema adoitan repetir a constatación da presenza relativamente escasa dese estrato na nosa área lingüística, en comparación con outras áreas da Península, e reiteran o breve repertorio de exemplos desa presenza no léxico e na toponimia coa indicación de que os máis deles chegaron a través do castelán e o portugués. Como a tradición filolóxica galega non conta con arabistas propiamente ditos, o que sabemos da historia dos arabismos galegos tamén nos chega en boa parte dos estudos e dos dicionarios etimolóxicos españois e portugueses (Asín Palacios, David Lopes, Cunha Serra, Machado, Corominas...), que en xeral ignoran calquera especificidade galega nos arabismos.

Frías Conde enumera algúns principios metodolóxicos que deben observarse no estudo do estrato árabe da nosa lingua. Un deles lembra a necesidade de explicar os arabismos galegos, e os hispánicos en xeral, non a partir do árabe clásico senón do dialecto hispanoárabe denominado *andalusí*, de onde o tomarían os mozárabes ou *dialectos romandalusís*, e destes os romances do norte. O autor ilustra ese principio co topónimo galego *Xurés*, portugués *Gerês*, partindo da idea de que está emparentado co fitónimo *jara* e é “da mesma orixe que Jerez en castelán” (Frías 2000, 164).

Non se pode defender a explicación arabista para o topónimo *Xurés / Gerês*. En primeiro lugar, porque non dá conta do artigo masculino singular presente nas denominacións galega e portuguesa, contrario a unha explicación que parte dun feminino plural. En segundo lugar, porque ignora a documentación antiga do topónimo. Isto con independencia da vaguidade con que se identifica co “Jerez en castelán”, pois se ben o granadino *Jeres del Marquesado* (paroxítono) admitiría un étimo relacionado co arabismo *xara, jara*, este non ten ningunha relación co gaditano *Jerez de la Frontera* (oxítono e de étimo escuro e controvertido) nin co estremeño *Jerez de los Caballeros* (tamén oxítono) documentado en fontes árabes medievais como *Xerixa* (Nieto Ballester 1997, 203).

Por outra banda, non temos constancia da existencia na área do *Xurés/Gerês* dunha denominación paralela ao castelán *jara* para as especies que noutras áreas peninsulares reciben ese nome ou formas emparentadas, e que deixan abundante

pegada na toponimia. Nesa área concreta, para o *Cistus ladanifer* e especies semellantes só rexistramos na toponimia formas emparentadas co común *chaguazo* (cast. *jaguarzo*).

2.2. Documentación medieval do topónimo

Edelmiro Bascuas recolle as atestacións medievais do topónimo no Tombo de Celanova, que confirman que o artigo actual é deglutinación dunha vogal velar inicial (Bascuas 2002, 238 e ss.). Extracto as seguintes, todas do século XI:

inter alpes montes Ezebrario iunias **Augeres** et mons Cabello discurrere rivulo Psallare [‘río Salas’] (ano 1024)

territorio Psallare sub alpes **Ugeres** (ano 1074)

et iacet subtus monte **Ugeres** discurrere rivulo Salar (ano 1096)

iacet in rio Caldo subtus **Ogeres** [...]et feret in termino de Lovios (ano 1100)

O *a-* inicial do documento máis antigo debe entenderse como aglutinación da preposición locativa (lat. *ad*).

Á vista das formas medievais, Bascuas interpreta a forma orixinaria do topónimo como un híbrido **Ugerense*, con radical hidrónimo prerromano *UG^w- e dous sufixos, un tamén prerromano *-er-* e o latino correspondente a *-ENSIS*, ao tempo que demostra que a redución *-ENSE* > *-és* estaba xa consumada nas datas dos documentos de Celanova e varios séculos antes (Bascuas 2002: 239). Segundo este autor, outros hidrotopónimos emparentados con este radical serían *Ogos* (lugar na parroquia e concello de Lobios, na mesma área do topónimo que nos ocupa), *Ogas* (lugar da parroquia de Cambeda, concello de Vimianzo) e posiblemente os diferentes *Oia* e tamén *Ames* (documentado como *Oames*, *Oiames* na Idade Media, como vimos).

2.3. Documentación anterior

Séculos antes das máis antigas atestacións celanovenses aducidas por Bascuas xa encontramos este topónimo en tres fontes clásicas que este investigador omite mais que serven para confirmar a súa hipótese.

A primeira é o documento chamado Itinerario de Antonino ou *Itinerarium provinciarum Antonini Augusti*, datado no século III (c. ano 280). Consérvanse del uns 20 manuscritos antigos, en copias de diferentes épocas, que citan máis de trescentas vías romanas, trinta e catro delas en Hispania, das cales se sitúan tres na Gallaecia. Na vía XVIII, de Bracara a Asturica, lese (entre parénteses as variantes segundo distintos manuscritos):

Salaniana XXI

Aquis Oreginis (Aquis Originis) XXI (XVIII)

Aquis Querquernis (Aquis Querquennis / Quercennis) XIII

(*Itinerarium Antonini*, 427-428; en Romero 1988: 113).

A segunda fonte é o texto xeográfico coñecido como Ravenate ou Anónimo de Rávena, composto arredor do ano 670 sobre documentación anterior, dos séculos IV e V. Na enumeración das localidades próximas a Braga (“item in ipsa Spania iuxta civitatem quam praediximus Augustam Bracaria dicitur civitas”) aparecen as tres *mansiones* do Itinerario de Antonino con variantes gráficas:

Salaniana

Aquis Ocerensis

Aquis cercenis

(*Ravenate*, 320; en Romero 1988: 151).

A segunda *mansio* das mencionadas tamén figura no mapa coñecido como Peutinger Karte ou Tabula Peutingeriana, atribuída a Castorius. Consérvase en copia feita nos séculos XII ou XIII dun modelo do VII, que á súa vez reproducía un orixinal do IV. Neste mapa figura coa mesma forma ca no anónimo de Rávena:

Aquis Ocerensis (Romero 1988: 114 e 155).

Moitos estudiosos da vía XVIII identificaron a *Aquis Oreginis / Originis / Ocerensis* destes textos co lugar actual dos Baños de Río Caldo, no concello de Lobios, ao pé dos montes do Xurés. A identificación vén supostamente confirmada polos achados arqueolóxicos, o mesmo que sucede coa *mansio* contigua *Aquis Querquennis* en Porto Quintela, hoxe no encoro das Conchas, no veciño concello de Bande. Tanto se a localización da primitiva *mansio* romana é ese lugar galego como se é a veciña localidade portuguesa de Caldas do Gerês, parece obrigado identificar o *Ogeres* medieval (seguramente oxítono) cos *Originis*, *Oreginis*, *Ocerensis* dos textos antigos, e todos eles co actual nome de O Xurés / O Gerês.

Deste xeito, das tres variantes rexistradas nas fontes clásicas, a máis fiel á forma que perviviu é o *Ocerensis* do Ravenate e do mapa de Peutinger, se ben as tres presentan como xorda a consoante velar sonora que deu orixe á palatal inicial actual, logo da deglutinación da vocal precedente. Non son necesarios moitos coñecementos de paleografía latina para saber que un *C* e un *G* poden confundirse doadamente na transmisión manuscrita, especialmente tratándose de palabras ou topónimos de linguas bárbaras.

No texto do Itinerario de Antonino ou na súa transmisión escrita produciuse talvez unha contaminación coa familia léxica latina de *origo*, *-inis*. Acontece adoito en situacións de contacto de linguas: os topónimos exóticos tenden a acomodarse ás estruturas fonéticas, morfolóxicas e léxicas da lingua que os incorpora. Entre os textos clásicos ilustrativos das dificultades que os topónimos indíxenas dos confíns do Imperio ofrecían aos latinofalantes, e que nos ensinan a ollar con cautela a toponimia das fontes documentais romanas, hai un fragmento de Plinio o Vello que precisamente se refire á área do Xurés ou Gerês actual:

Simili modo Bracarum XXIII civitates capitum, ex quibus praeter CCLXXXV ipsos Bracaros Bibali, Coelerni, Callaeci, Equaesii, Limici, Querquerni citra fastidium nominentur.

(Plinio, *Natural Historia*, III, 28; Romero 1988: 65).

Entre as vinte e catro *civitates* do territorio dos Brácaros, que sumaban 285000 habitantes, Plinio fai constar unicamente sete cuxos nomes poden escribirse ou pronunciarse á latina sen moito problema (*citra fastidium nominentur*). Entre as deza-sete restantes, omitidas por causa dos nomes problemáticos, acharíase verosimilmente a mesma que serviu para denominar unha das *mansiones* do Itinerario de Antonino e un territorio tan extenso como o Xurés / Gerês actual.

En conclusión, as variantes *Aquis Originis*, *Oreginis* e *Ocerensis* dos clásicos constitúen deturpación de **Aquis Ogerensis*. A correspondente forma de acusativo singular, **Ogerense(m)* é case coincidente co **Ugerense* proposto por Bascuas como étimo das formas medievas *Ogeres*, *Ugeres* e as formas actuais. A diferenza da vogal velar inicial (u/o) é irrelevante para a súa deglutinación (medieval ou posterior) ao ser interpretada como artigo masculino, mais talvez non o sexa tanto á hora de sinalar para o étimo unha raíz hidronímica prerromana en U- cando as fontes latinas rexistran o-. En calquera caso, temos que aceptar que se trata dun topónimo prerromano, polo menos no seu radical. O sufixo podería ser efectivamente o latino -ENSE, como defende Bascuas. Elixio Rivas, non obstante, referíndose a topónimos ourensáns como *Brués*, *Palmés* ou *Bousés*, sostén a existencia dun sufixo indíxena foneticamente equivalente que supostamente se asimilou ao -ENSE latino (Rivas 1989: 110).

2.4. O Gerês e O Xurés

A evolución desde **Ogerense* ás denominacións actuais ofrece resultados regulares en todos os aspectos. O paso de -ENSE a -ES ten testemuños hispánicos que remontan ao século VI (Menéndez Pidal 1950: § 48). A palatalización e fricativización da consoante velar oclusiva ante vogal palatal é tamén regular, co resultado sonoro conservado en portugués e enordecido no galego. Tamén é máis conser-

vadora a forma portuguesa no relativo ao vocalismo, pois no galego operou unha disimilación da vogal pretónica.

Nas dúas variantes produciuse unha aférese ou deglutinación da vocal inicial interpretada como artigo. É este un fenómeno relativamente frecuente en galego e portugués, tanto na toponimia coma na lingua común (*o bispo* < *obispo* < EPISCOPU), e ten paralelos noutras áreas lingüísticas (cast. *atril* < LECTORILE, cat. *Igalada* < *L'Aigua Lada* < AQUA LATA). Na toponimia galega son abundantes os exemplos de deglutinación semellante. Por citar só topónimos ben coñecidos: A Baña (*Avania* na documentación antiga), O Incio (*Onitio*), Ríos (*Orreolos*) e a forma A Maía (*Amaea*). Con eles podemos contar os diferentes O Grove (*Ocobre*) da xeografía galega e portuguesa¹², onde tamén houbo interpretación da vogal inicial como artigo singular. Un exemplo do plural son os ourensáns As Maos / As Maus e o portugués Mãos (concello de Ponte de Lima), que son topónimos medievais que conteñen un xentilicio (*Asmanos*) derivado do hidrónimo Asma, por alusión a colonos medievais procedentes da comarca de Chantada.

Referencias bibliográficas

- Asín Palacios, M. (1940): *Contribución a la toponimia árabe de España* (Madrid: CSIC).
- Baión, C. de (1976) [pseudónimo de X. M^a Álvarez Blázquez]: *Os nomes da terra recollidos do pobo* (Vigo: Castrelos).
- Baliñas Pérez, C. (1998): *Gallegos del año mil* (A Coruña: Fundación P. Barrié de la Maza).
- Bascuas, E. (2002): *Estudios de hidronimia paleoeuropea gallega*. Anexo 45 de *Verba* (Santiago: Universidade de Santiago de Compostela).
- Blanco, L. (1995): *Estudio de toponimia menor en la parroquia de Santa María de Samieira (Poyo - Pontevedra)* (Pontevedra: Diputación Provincial).
- Carré Aldao, E. (1980) [reimp. facs. da ed. de 1936]: *La Coruña*, 2 vols., en Carreras Candi, F. (1928-1936): *Geografía general del Reino de Galicia* (A Coruña: Ediciones Gallegas).
- Corominas DCECH = Corominas, J. / Pascual, J. A. (1980-1991): *Diccionario crítico etimológico castellano e hispánico* (Madrid: Gredos).

¹² Por exemplo a localidade portuguesa do Grove (Fontoura, Valença). Figura sen artigo no *Dicionário Corográfico de Portugal* de Frazão. Nos países de lingua portuguesa tradicionalmente non se rexistra o artigo dos topónimos na rotulación viaria e nos nomenclátors, de xeito que de non coñecer os lugares, un non sabería que teñen artigo Porto, Estoril, Caldas da Rainha, Rio de Janeiro, Recife ou Bahía.

- Cortese, A. (1912): *Onomástico Medieval Portugués* (Lisboa: Imprensa Nacional).
- DRAE = Real Academia Española, *Diccionario de la lengua española*, Vigésima segunda edición, 2001 (Madrid: Real Academia Española). Cito pola edición en rede: <http://buscon.rae.es/diccionario/drae.htm>.
- Díez Sanjurjo, M. (1906) = M.D.S. [Manuel Díez Sanjurjo]: “Documentos históricos. Allariz”, *Boletín de la Comisión Provincial de Monumentos de Orense*, 3: 95-97 (Ourense: Comisión Provincial de Monumentos).
- Ferreiro, M. (1995): *Gramática histórica galega* (Santiago: Laivento).
- Frazão, A. C. do Amaral (1981): *Novo Dicionário Corográfico de Portugal* (Porto: Ed. Domingos Barreira).
- Frías Conde, X. (2000): “O elemento árabe en galego”, *Revista Galega de Filoloxía*, 1: 157-171.
- González Garcés, M. (1987): *Historia de La Coruña* (A Coruña: Caixa Galicia).
- Houaiss = Instituto Antônio Houaiss (2001): *Dicionário Eletrônico da Língua Portuguesa* (Rio de Janeiro: Editora Ojetiva Ltda.).
- Hoz, J. J. de (1963): “Hidronimia antigua europea en la Península Ibérica”, *Emérita*, 31: 227-242.
- Lopes, D. (1968): *Nomes Árabes de Terras Portuguesas* (Lisboa: Sociedade de Língua Portuguesa / Círculo David Lopes).
- Lorenzo, R. (1977): *La traducción gallega de la crónica General y de la Crónica de Castilla*. Vol. 2. *Glosario* (Ourense: Instituto de Estudios Orensanos ‘Padre Feijoo’).
- Madoz, P. (1845): *Diccionario Geográfico, Estadístico e Histórico de España y sus Posesiones de Ultramar*, s.v. *Almeiras*. Cito pola ed. facsimilar de 1986 (Santiago: Libros Galicia, Edit. Breogán).
- Menéndez Pidal, R (1950 [3ª ed]): *Orígenes del español. Estado lingüístico de la Península Ibérica hasta el siglo XI* (Madrid: Espasa-Calpe).
- Navaza, G. (2003): “Sarmiento toponimista”, en *Actas do I Congreso Internacional de Ciencias Onomásticas Frei Martín Sarmiento*, Santiago de Compostela, 5-7 de setembro de 2002 (no prelo).
- Nieto Ballester, E. (1997): *Breve diccionario de topónimos españoles* (Madrid: Alianza Editorial).
- Millán González-Pardo, I. (1987): *Toponimia del concejo de Pontedeume y cartas reales de su puebla y alfoz* (A Coruña: Deputación Prvincial).

- Pensado Tomé, J. L. (ed.) (1973): Fr. Martin Sarmiento: *Catálogo de voces y frases de la lengua gallega* (Salamanca: Universidad de Salamanca).
- Pensado Tomé, J. L. (ed.) (1975): Fr. Martin Sarmiento: *Viaje a Galicia (1745)* (Pontevedra: Museo de Pontevedra).
- Porto Editora (2002): *Dicionário da Língua Portuguesa Online* (<http://www.portoeditora.pt/dol/>).
- Rivas Quintas, E. (1989): *Lingua galega. Historia e fenomenoloxía* (Lugo: Alvarellos).
- Rivas Quintas, E. (1994): *Lingua galega. Nivéis primitivos* (A Coruña: Laiovento).
- Romero Masiá, A. M. / Pose Mesura, X. M. (1988): *Galicia nos textos clásicos* (A Coruña: Museo Arqueolóxico da Coruña).
- Santamarina, A. (2001) [2ª ed.]: *Diccionario de diccionarios* (Santiago: Fundación Barrié de la Maza)
- Veiga Arias, A. (1983): *Algunas calas en los orígenes del gallego* (Vigo: Galaxia).
- Villar Liébana, F (2000): *Indoeuropeos y no indoeuropeos en la España preromana* (Salamanca: Universidad de Salamanca).
- Viterbo (1993) = Fiúza, M (ed. crit) [2ª reimp.]: Fr. J. de Santa Rosa de Viterbo: *Elucidário das palavras, termos e frases que em Portugal antigamente se usaram e hoje regularmente se ignoram (1798-99)* (Porto-Lisboa: Livraria Civilização).

Fontes documentais:

- Almeida = Iglesias Almeida, E. (1984): *Los antiguos "portos" de Tuy y las barcas de pasaje a Portugal* (Tui: Museo y Archivo Histórico Diocesano).
- Celanova = Andrade, J. M. (1995): *O Tombo de Celanova*, 2 vols. (Santiago: Consello da Cultura Galega).
- Devanceiros = Ferro Couselo, X. (1967): *A vida e a fala dos devanceiros*, vol. I (Vigo: Galaxia).
- Espadanedo = Pereira Ferreiro, S. (1979): *El monasterio de Santa María de Xunqueira de Espadanedo*. Anexo IX de *Boletín Auriense* (Ourense: Museo Arqueolóxico Provincial).
- FDU = Justo, M. J. / Lucas Álvarez, M. (1991): *Fontes documentais da Universidade de Santiago de Compostela* (Santiago: Consello da Cultura Galega).

- Maia = Maia, C. de Azevedo (1986): *História do Galego-Português. Estado lingüístico da Galiza e do Noroeste de Portugal desde o século XIII ao século XVI (com referência à situação do galego moderno)*, (Coimbra: Instituto Nacional de Investigação Científica).
- Oseira = Romaní, M. (1990 [vols. I e II], 1993 [vol. III]): *Colección diplomática do mosteiro cisterciense de Santa María de Oseira* (Santiago: Tórculo).
- Ramirás = Lucas Álvarez, M. / Lucas Domínguez, P. P. (1988): *San Pedro de Ramirás. Un monasterio femenino en la Edad Media* (Santiago: Caixa Galicia).
- Ribadavia = Enríquez Paradela, M. C. (1987): *El monasterio y convento de Santo Domingo de Ribadavia. Colección diplomática. Anexo VIII de Boletín Auriense* (Ourense: Museo Arqueolóxico Provincial).
- Rocas = Duro Peña, Emilio (1972): *El monasterio de San Pedro de Rocas y su colección documental* (Ourense: Instituto de Estudios Orensanos 'Padre Feijoo').

E ja quites son los Mans. Nota paleográfica sobre o v. 4 da cantiga B 454 de D. Garcia Mendiz d'Eixo

Xosé Bieito Arias Freixedo
Universidade de Vigo

Son de todos coñecidas as dificultades que presenta para os editores a cantiga B 454 de D. Garcia Mendiz d'Eixo, non tanto polo feito de que estea escrita en provenzal, ou de que en aparencia tente reproducir esta lingua, senón porque probablemente por esta razón nalgún momento da súa transmisión manuscrita un copista tivo dificultades para interpretala e a versión que nos foi transmitida chegounos deturpada e presenta grandes problemas de lectura e de interpretación. Abonda con remitir ás varias edicións para comprobar que as interpretacións son totalmente diverxentes. Por exemplo, a dúas edicións realizadas con detemento como son a de J. Marie D'Heur (1973: 93-104) –demasiado afastada do texto manuscrito– e a recente edición de Graça Videira Lopes (nº 12), das que reproducimos a primeira estrofa, onde se atopa o noso verso, o que nos serve para situalo en contexto e ademais para exemplificar a considerábel diverxencia entre ambas versións.

A la mazo on e[s] la corona,
e li parent son tan,
e la terra es tro bona,
e ja quyte sson li chan,
c'ara me volh tornar
a Sousa, a lo mon logar
que m'adola e ma sauda
(D'Heur, p. 100)

Alá nazque la Torona
e los pavous son tan[s]
e la terra é tro bona!
Eia! quites son los chans!
C'ora me volho tornar
a Sousa, a lo mon logar,
que me adota e me saudona.
(Lopes, p. 34)

Non imos realizar aquí unha edición de todo o texto; simplemente chamaremos a atención sobre unha deficiente lectura do verso cuarto da composición, que se repite de editor en editor, e que coidamos se pode corrixir se se analizan as características paleográficas do manuscrito un pouco pormenorizadamente.

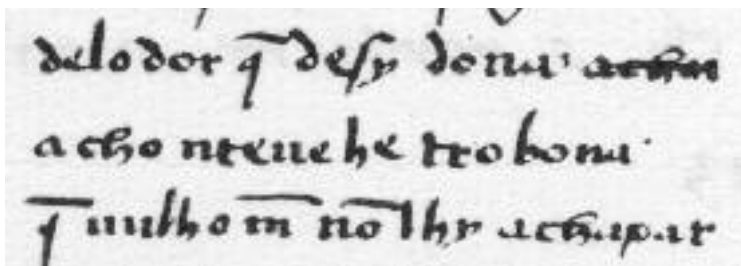
D’Heur le este verso no manuscrito do seguinte xeito: *Eia. quytes fõ los chaus*

G. V. Lopes le tamén como *chaus* a última palabra do verso, aínda que edita *chans* e propón interpretala como ‘cantos’ ou talvez como ‘prantos’. Aínda na sección final de notas (p. 553) propón outra interpretación posíbel mantendo a lección do manuscrito, que ela le *chaus*, co sentido de ‘calores’: “(acabaram-se os calores)”.

É sobre este último termo que queremos chamar a atención, pois no manuscrito non se le *chaus*, senón *Maus*, como trataremos de demostrar, cotexando as grafías con outras grafías da mesma man, do mesmo folio e da mesma cantiga.

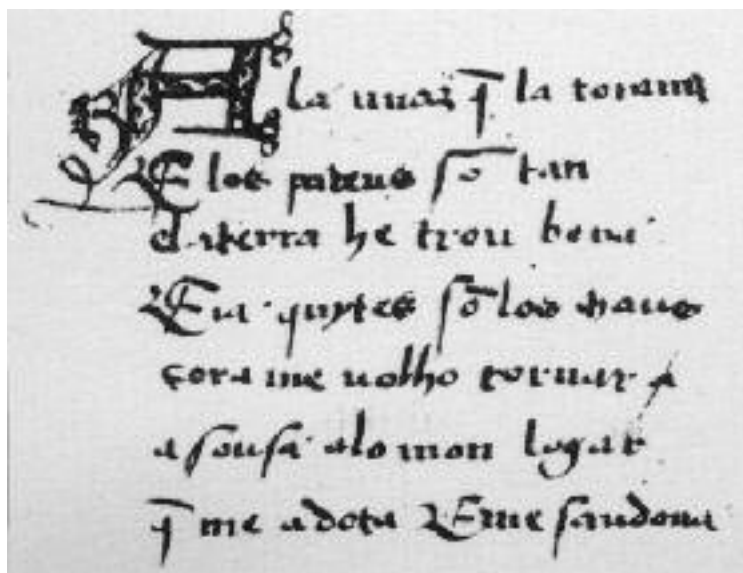
O cotexo da grafía inicial deste termo con outras grafías que aparecen na mesma cantiga, copiadas pola mesma man, para representar o fonema africado palatal xordo, o dígrafo *ch*, permítenos concluír que neste verso cuarto estamos perante unha grafía diferente. Na columna b do folio 99 (p. 225 da ed. facsímile), onde continúa a copia da cantiga, rexístrase este dígrafo en tres ocasións, e nas tres presenta o mesmo trazado: o *c* unido ao *h* sen erguer o trazo, de xeito que o *h* presenta no hástil o lazo característico da cursividade da escrita.

Os tres exemplos aparecen en liñas seguidas, o primeiro deles riscado, porque se debe a un salto de liña do copista, como pode comprobarse na seguinte imaxe:



Na última palabra das liñas primeira e terceira temos o dígrafo *ch* (no primeiro caso o copista riscou logo a palabra ao decatarse de que pertencía ao verso seguinte). Despois do *a* inicial da segunda liña volve aparecer este dígrafo.

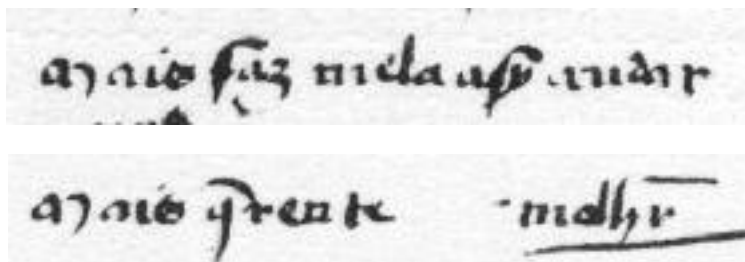
Se o compararmos agora coa grafía que se rexistra na palabra final do cuarto verso da primeira estrofa, veremos que estamos perante outra grafía diferente (ofrecemos a imaxe da estrofa íntegra para que o lector ou lectora poida cotexar con ela as edicións críticas ofrecidas máis arriba):



Parece claro que a grafía non é a mesma que a da outra imaxe. É importante ter en conta o *usus scribendi* do amanuense, que parece indicar que o hábito do copista era o de escribir o dígrafo *ch* sen erguer o trazo e con bucle ou lazo no hástil do *h*. En moi última instancia o *h* podería selo, mais o trazo que o precedería en ningún caso podería corresponder a un *c*.

Cal é a grafía do v. 4, pois?

O certo é que a aplicación dunhas nocións básicas de paleografía permítenos propor a hipótese de que nos atopamos perante unha das variantes do *M* maiúsculo que presenta a gótica cursiva. É mais, no mesmo folio 99, o mesmo copista emprega unha grafía similar como inicial en dous versos da cantiga seguinte, aos que corresponden as seguintes imaxes:



Polo tanto haberá que ler o rimante do noso verso como *Maus*.

Mais a interpretación do termo *maus* como forma plural substantivizada do adxectivo *mau* (mao), derivado de MALUS, que é a que xorde a primeira vista e que concordaría co sentido do conxunto da cantiga (o poeta, que está exilado, ve chegada a hora de tornar á súa patria, porque os *maus*, ou sexa, os ‘maos’, os que llo impedirían, agora *son quites*, isto é, desapareceron, ou foron desprazados do poder), non é a correcta. En efecto, se tivermos en conta o esquema de rimas da cantiga, vemos que é o mesmo na primeira e segunda estrofas: *ababcca*. Na segunda estrofa a rima *b* é *-al*, e nesta primeira estrofa é preciso corrixir no v. 2 (*tan[s]*) engadindo un *-s* para emparellala coa rima do v. 4, en que á súa vez hai que corrixir *u* por *n*, para que a rima *b* sexa *-ans* nos dous versos.

Atopámonos así co rimante *Mans*. Mais, cal é o valor desta palabra no texto?

Non resulta doado decantarse por un significado concreto, xa que este varía dependendo do étimo a que se remonte o termo. Tendo en conta o sentido que dun xeito global parece percibirse no texto, relativo ao desexo do protagonista de tornar ao seu lugar patrio, e considerando as circunstancias históricas reais do autor, que viviu en persoa a experiencia do desterro, coidamos que o termo provenzal *mans* podería remontarse ao baixo latín MANSUS, termo empregado abundantemente na literatura documental medieval en latín (cf. Du Cange, *Glossarium Mediae et infimae latinitatis*, s. v.) cun valor equivalente ao de *villa*, ou sexa, ‘casa de campo, coas construcións adxacentes e terreos de cultivo’, ou tamén equivalente a *vicus*, isto é, ‘lugar, aldea’.

Este valor, que se podería xeneralizar como ‘propiedades, posesións’, cadra perfectamente co sentido xeral da composición, do que se desprende o desexo do protagonista de ‘*tornar / a Sousa*’, ao seu lugar familiar, agora que as súas posesións, os seus *mansos* (*mans*) xa estaban libres (*quites*).

Aínda sería posíbel outra interpretación do termo *mans* se o entendésemos como un substantivo derivado regresivo de *mandar* (<MANDARE), co sentido de ‘ordes,

mandatos'. Sería o mesmo significado que se rexistra na cansó do trobador provenzal Marcabré *A la fontana del vergier* (vid. Riquer, nº 20, v. 27): *que fai los mans e los prezicx*: “que dispón as ordes e as predicacións”.

No noso texto este significado tamén tería sentido se o termo *mans* se referise ás ordes que obrigaban ao protagonista a estar lonxe da súa terra (*Sousa*), á que agora quere tornar, porque as tales ordes ou disposicións xa foron revocadas: *e ja quites son los mans*.

Coidamos, en definitiva, que calquera das dúas interpretacións aquí propostas melloran as existentes. En todo o caso, son máis xustificábeis desde o punto de vista paleográfico.

Referencias Bibliográficas:

- CANCIONEIRO DA BIBLIOTECA NACIONAL (Colocci-Brancuti) Cod. 10991, I. (1992). Reprodução facsimilada. Edição de L. f. Lindley Cintra (Lisboa: Imprensa Nacional / Casa da Moeda / Biblioteca Nacional).
- HEUR, J. M. d' (1973): *Troubadours d'oc et troubadours galiciens-portugais. Recherches sur quelques échanges dans la littérature de l'Europe au Moyen Age* (París: Fundação Calouste Gulbenkian).
- LOPES, G. Videira (1994): *A sátira nos cancioneiros medievais galego-portugueses* (Lisboa: Imprensa Universitaria / Editorial Estampa).
- RIQUER, M. de (1993): *Los trovadores. Historia literaria y textos* (Barcelona: Ariel).

A lexitimidade da adaptación como estratexia tradutiva

Robert Neal Baxter
Universidade de Vigo

Resumo:

Encadrado dentro do contexto do eterno debate –vello mais actual a un tempo– que se desenvolve no ámbito dos estudos da tradución respecto da lexitimidade da adaptación en canto estratexia tradutiva, apoiado por unha serie de exemplos prácticos ilustrativos, sempre desde a perspectiva da tradución global mediante equivalencias dinámicas dentro do cadro da transferencia e intercambios intersistémicos, este artigo sostén que a adaptación non é non só lexítima senón que necesaria e inclusive inevitábel desde o momento en que se muda dunha lingua para unha outra. Advoga, así mesmo, para a plena asunción desta realidade por parte do tradutor para se poder aproveitar de xeito creativo, asegurando así a plena funcionalidade do novo texto (traducido) dentro do seu sistema de recepción.

Palabras chave:

Aadaptación. Fidelidade. Funcionalidade. Transferencia intersistémica.

Abstract:

Set within the framework of the eternal and ongoing debate raging within translation studies regarding the legitimacy or not of adaptation as a basis translation strategy, backed by concrete examples, this article set within the context of global translation via dynamic equivalences based on intersystemic transfer and exchange argues that adaptation is not only legitimate, but indeed necessary and unavoidable when language switching is involved. It is argued that by becoming fully aware of this fact, the translator could use it creatively to his advantage, thus ensuring the full functionality of the new (translated) text within its receptor system.

Key words:

Adaptation. Faithfulness. Functionality . Intersystemic transfer.

Ao longo destas liñas, propomos reflexionar sobre un elemento central, non isenta de polémica, que subxace en todo o momento á focaxe pola que advogamos respecto da necesidade de elaborar unha tradución global e plenamente funcional e orientada cara á aceptabilidade do texto traducido final como un texto plenamente funcional e autónomo dentro do polisistema literario de chegada. Trátase do debate sobre se o tradutor pode, ou inclusive debe, adaptar o texto de partida e, por conseguinte, se a adaptación é unha estratexia lexítima, ou se, pola contra, sitúa quen

a práctica fóra do fenómeno que se dá en denominar ‘tradución’. Noutras palabras, débese considerar a ‘adaptación’ como unha estratexia tradutiva máis ou débese entender como unha actividade á parte que, de se querer inserir dentro do dominio da tradución e gozar desta denominación, se constituiría, de facto, nunha ‘traizón’ cara ao texto orixinal.

Rabadán (1991) por exemplo, partindo dunha concepción, que compartimos, da tradución entendida como un acto de transferencia intersistémica mediante o establecemento de equivalencias dinámicas para acadar un texto autónomo e funcional dentro do polisistema de valores de chegada, fala en termos non de inserir o texto traducido dentro do polisistema de chegada, senón da criación dun texto novo dentro do sistema de chegada. Unha visión da tradución entendida nestes termos apaga case totalmente as fronteiras que delimitan tradicionalmente a ‘tradución’ da ‘adaptación’, onde o tradutor non é só unha especie de interface lingüístico-cultural que pasa elementos do texto orixinal a través dun filtro para xerar un texto substitutivo noutra lingua, senón un auténtico suxeito activo e creativo que ha de asumir plenamente as responsabilidades que decorren do seu papel como ‘autor segundo’ encargado de reescribir un texto novo a partir doutro orixinal.

Existen, de facto, ámbitos determinados onde a adaptación motivada pola procura de reproducir un efecto moi determinado no público receptor do texto similar ou equivalente ao efecto producido polo texto orixinal no seu público receptor orixinal de referencia constitúe non só unha posibilidade tradutiva, senón que é a única alternativa, sen que deixe de ser, por tanto, tradución. Tal é o caso, por exemplo, da publicidade:

We should point out that equivalent effect is not always the primary concern of advertisement translating, and at times cultural transposition or adaptation is inevitable in the translating of advertisement. This is not only legitimate, but also necessary, in the attainment of the *skopos* (Kim-lung 1999: 104).

Mais a adaptación non se restrinxe como estratexia tradutiva lexítima (e necesaria) ao ámbito da publicidade ou en casos especializados como a denominada ‘localización’¹, senón que resulta ser un alternativa posíbel en case todos e calquera ámbito, excluíndo algúns casos moi marxinais que comentaremos a continuación. Mais con todo, aínda que o grao de adaptación efectuada variará dun texto para un outro, dun tradutor para un outro e dun momento e un lugar concretos para un momento e un lugar diferentes, a adaptación sempre estará presente como un elemento esencial e básico de calquera estratexia tradutiva. O esencial consiste en recoñecer este

¹ ‘Localización’ é o termo que se utiliza habitualmente para se referir á tradución de *software* para se tornar totalmente aceptábel para o sistema receptor, sendo así un tipo de adaptación tradutiva tematicamente determinada.

facto, asumilo e aproveitar o potencial que ofrece o uso consciente da adaptación como estratexia tradutiva.

Para ilustramos o facto de a tradución, pola súa propia natureza, sempre levar implícito un transfondo básico e ineludíbel de adaptación xa polo simple facto de implicar unha mudanza de linguas, pasaremos de contado a analisarmos as mudanzas efectuadas ao traducir un texto chinés para o español que decorren das mudanzas necesarias ao pasar dun sistema lingüístico-cultural determinado para un outro cuxa estrutura dista desta, pois se non se diferenciase, sería, loxicamente, a mesma lingua e non faría falta tradución.

En definitiva, traducir implica reformulación, e calquera reformulación leva implícita unha mudanza máis ou menos importante, ao seren indisolubelmente ligadas a forma e a función (lingüísticas ou non):

Reformulating something transforms the ways in which meanings may be made with it, and in this sense, form and content are inseparable (Chandler 1995: 16).

Escollemos para estes efectos ilustrativos dous sistemas lingüísticos estruturalmente moi diferentes a todos os niveis (fonolóxico, morfosintáctico...) así como culturalmente (aquí, por exemplo, abordaremos a cuestión da poesía tal e como se entende dentro do polisistema literario chinés), para pormos máis facilmente en destaque a natureza e as implicacións das mudanzas realizadas: o chinés clásico (a forma utilizada para a poesía clásica en cuestión) é unha lingua monosilábica onde existe unha relación directa entre o grafema, a sílaba e o morfema (que prevalece sobre o concepto occidental de 'palabra'), que non distingue tan nitidamente entre as clases gramaticais (adxectivo e verbo forman unha única clase formalmente), non coñece a distinción temporal pasado-presente-futuro mais si fai fincapé na distinción de aspecto perfectivo-imperfectivo, non ten nin xénero nin número, sen artigos determinados ou indeterminados, sen flexión morfolóxica... En definitiva, unha lingua radicalmente diferente, canto á estrutura, do español, lingua para que se ha traducir.

Ora ben, débese ter presente que aínda que as mudanzas entre sistemas máis aparentados lingüisticamente foren menos perceptíbeis á primeira vista, sempre estarán presentes a un grao ou a un outro, por exemplo a gramaticalización da noción de aspecto entre o español (*estaba comiendo - comía - comí - he comido*) e o galego (*estaba a comer - comía - comín*).

Enumeraremos algunhas das mudanzas efectuadas ao se compararen un extracto dunha poesía chinesa e a súa tradución, neste caso para o español. Trátase dunha poesía composta polo o Presidente Mao Zedong en 1958 seguindo a métrica dun

tipo de poesía clásica denominada *qiliu*. Seguiremos as explicacións que nos dan Moravia / Mancuso (1975: 41-44):

红雨随心翻作梁	<i>hóng yǔ / suí xīn / fān zuò liáng</i> <i>vemello chuvia / seguir corazón / cambiar facer</i> <i>ponte</i>
青山着意化为桥	<i>qīng shān / zhāo yì / huā qiáo</i> <i>verde montaña / mandar vontade / tornarse</i> <i>facer ponte</i>
天连五岭银锄落	<i>tiān lián / wǔ líng / yín chú lào</i> <i>ceu unir / cinco cordilleiras / prata sacho caer</i>
地动三河铁臂摇	<i>dì dòng / sān hé / tiě bì yáo</i> <i>terra move / tres río / ferro brazo sacudir</i>

Eis a tradución que nos ofrecen Moravia / Mancuso (1975: 129):

Lluvia rosa de flores de melocotón
que a un ademán de aqueles se transforman en río,
Verdes montes que a su orden
se transforman en puentes.
Sobre las Cinco Cadenas que llegan al cielo
se abaten azadas de plata,
brazos de hierro desvían los Tres Ríos,
transforman toda la tierra

O noso propósito aquí non é o de avaliarmos a tradución que se nos propón, senón, simplemente, recalcarmos o número e a importancia das mudanzas inevitábeis que se realizan a todos os niveis ao pasar do orixinal á tradución.

- ++ / ** / ++ * **R** Eis como funciona unha das dúas posíbeis métricas e a rimas para os poemas do tipo *qiliu*.
- ** / ++ / * * + **R** indica a rima.
- ++ / ** / ++ * **R** / indica un contrapunto (gramatical e semántico)
- ++ / ** / * * + * **R** * indica o ton ‘baixo’ (na transcripción *pinyin* este ton está sinalado polo uso do macron, ex. <ā>)
- ** / ++ / + * * **R** + indica os outros tons chamados ‘oblicuos’
- ** / ++ / * * + * **R** A poesía de Mao Zedong de que se trata aquí segue esta métrica, aínda que non con absoluta fidelidade.
- ++ / ** / ++ * **R** (véxase Moravia / Mancuso 1975: 41-46)

Así, fica evidente, en primeiro lugar, que o orixinal é moito máis compacto que a tradución, que conta con case dúas veces máis morfemas/palabras (27:50) e máis de tres veces máis sílabas (27:89). Até hai dúas veces máis liñas na poesía traducida (8) do que na orixinal (4). Visualmente, na súa versión orixinal en chinés, a poesía forma ringleiras ordenadas e simétricas, o que deixa de ser o caso en español coas súas liñas irregulares. Iso débese tanto á propia natureza lingüística da lingua chinesa clásica, onde existe unha relación directa entre palabra-morfema-sílaba, e tamén á natureza logográfica da escritura do chinés (véxase Sampson 1985: 145), que pasa a ser, na tradución, unha escritura alfabética.

Alén de non manter as rimas finais que fan parte da métrica ríxida da forma poética do orixinal e do que non discutiremos a lexitimidade tradutiva, o que si se perde irremediabelmente e contra a vontade de elección do propio tradutor é o sistema interno de rimas de tons que non se pode plasmar en español por non ser esta unha lingua de tons. Alén de ser un mero elemento decorativo, a métrica tónica e a propia estrutura do orixinal tamén determinan a análise lingüística no momento de traducir:

Vemos así como las reglas prosódicas se convierten en realidad en auténticas reglas gramaticales, que cuando son respectadas [...], contribuyen a definir el sentido de un verso, a precisar si una palabra determinada se interpreta como adjetivo, sustantivo, etc. Nos hallamos, pues, una vez más, ante reglas posicionales, si bien de tipo distinto de las que regulan la prosa. (Moravia / Mancuso 1975: 43).

Ora ben, o facto de se realizaren tantos e tan profundos cambios non fai que a tradución sexa nin imposible –de facto, existe– nin tampouco que non sexa unha tradución de má cualidade. O que si nos ensina é que a adaptación non é só factíbel e lexítima ao se traducir un texto, ao tentarmos efectuar eficazmente unha transferencia intersistémica, senón que é absolutamente inevitábel. A adaptación é, noutras palabras, a esencia mesma da tradución.

Para axudar a entender até que punto resulta imprescindible ficar consciente das posibilidades de mudanzas que se teñen que procurar no momento de funcionarmos como auténticos mediadores interculturais, propomos considerar o caso da tradución de frases feitas, sobre todo aquelas que conteñen algunha que outra imaxe visual evidente:

Língua	Frase feita	Tradución literal
inglés	it's raining cats and dogs	'chove cans e gatos'
galés	mae hi'n bwrw hen wragedd a ffyn	'lánzanse vellas e paos'
francés	il pleut des cordes	'chove cordas'
catalán	plou a bots i barrals	'chove chicados e bocois'
galego	chove a deus dar	–

A menos que se tratar dunha tradución cun claro propósito burlesco, para se manter 'fíel ao orixinal', a ningún tradutor de catalán, poñamos por caso, se lle pasaría pola cabeza traducir, *it's raining cats and dogs* por *plou a gats i gossos*. Automaticamente, sen pensalo dúas veces, a frase en inglés pasaría a ser *plou a bots i barrals*. Múdase a frase feita por outra equivalente no sistema de recepción, mais ao mesmo tempo fica alterada unha imaxe –non é o mesmo, claro é, gatos e cans que chicados e bocois...

En definitiva, fica patente, dun lado, que canta máis distancia existe entre os sistemas de partida e de chegada –nese noso primeiro caso entre o chinés e o español respectivamente– máis clara se torna a inevitabilidade de se realizaren cambios máis ou menos importantes ao pasar dun sistema para un outro.

Segundo a hipótese que propuxeron no seu día Sapir e Whorf, existiría unha relación de influencia mutua entre as estruturas lingüísticas propias e características de cada lingua humana e a percepción que se teña da realidade que nos rodea. Isto é, que os seres humanos utilizamos a linguaxe para exprimir o mundo ao tempo que a clasificación que se precisa facer do mundo para se poder exprimir mediante a linguaxe implica que esta mesma linguaxe, unha vez interiorizada como o noso principal medio de comunicación, axe como un filtro que determina como percibamos o mundo segundo as estruturas que permite cada linguaxe, de xeito que cada falante dunha lingua diferente percibiría o mundo dun xeito diferente, ou cando menos así o exprimirá.

Así pois, mudar de lingua leva necesariamente a unha mudanza na percepción e representación que se fai da realidade, sendo esta unha mudanza inelutábel. Desde xeito, pois, enténdese como, sob o noso punto de vista, as mudanzas lingüísticas levan implícitas mudanzas correspondentes e inevitábeis na substancia da mensaxe en maior ou menor grao, tendente sempre cara á adaptación respecto do pólo receptor.

Esta tese que defenderon no seu tempo Sapir e Whorf non é isenta de críticas e de polémica, mais, sen entrarmos nos pormenores do debates lingüísticos correspon-

dentés, achamos que é interesante para os fins que nos propomos, ou sexa, indagarmos na natureza da adaptación extralingüística como un corolario inevitábel de calquera mudanza dunha lingua para unha outra, de tal xeito que a adaptación sempre forme parte da tradución e non se trate dunha disciplina diferente, nin moito menos dunha ‘traizón’ cara ao orixinal.

Achamos precipitado, pois, rexeitar a hipótese Sapir-Whorf tal e como pretende Pinker (1994: 57), sempre e cando se evitar o ‘determinismo lingüístico radical’ (*strong determinism*) que ás veces vén asociado a este termo, que consiste na crenza de que o mundo exterior se percibe a través do filtro da linguaxe sen máis:

Whilst few linguists would accept the Sapir-Whorf hypothesis in its ‘strong’, extreme or deterministic form, many now accept a ‘weak’, more moderate, or limited Whorfianism, namely that the ways in which we see the world may be *influenced* by the kind of language we use (Chandler 1995: 17).

Advogamos aquí, pola contra, por un ‘determinismo relativo’ (*weak determinism*) en que os procesos cognitivos e o pensamento humano estarían influenciados pola linguaxe que utilizamos sen, por tanto, que a linguaxe chegue a determinar como se percibe o mundo:

Weak determinism, however, holds that thought is merely affected or influenced by our language, whatever that language may be. This version of determinism is widely accepted today (Campbell 1997: 5).

Partindo desta hipótese, onde as formas e as estruturas das linguas que se utilizan para exprimir a realidade (inclusive sentimental, emotiva...) condicionarian, en certa medida a maneira en que se pode exprimir tal realidade ao influír na súa percepción, algúns comentaristas chegan á conclusión de que a tradución é xa unha tarefa imposíbel ou cando menos extremadamente complexa debido a que, loxicamente, ningunha lingua partilla plenamente as estruturas dunha outra e, por tanto, calquera mudanza implicaría unha falta de ‘fidelidade’ para co orixinal. Outros (Campbell 1997: 16) suxiren, inclusive, que calquera reformulación dentro dunha mesma lingua levará implícitas mudanzas no sentido orixinal.

Ora ben, iso só é válido desde o momento en que se considere a tradución como un proceso que pretende reproducir nun sistema meta a través dunha mudanza de lingua todos e exactamente os elementos semánticos presentes no orixinal. Mais esta concepción parte de entrada dunha visión da tradución como un exercicio orientado cara ao pólo de partida e cara ao de chegada. Neste sentido, coincidiremos en que a tradución é por si imposíbel. Ora ben, a tradución existe, e por tanto ha de ser posíbel. Só é posíbel ao refocala correctamente, aceptándoa polo que en realidade é: unha aproximación a partir dun texto orixinal interpretábel desde o polisistema

de valores de partida para unha lingua diferente e un outro polisistema de valores. Nestes termos, pois, enténdese como a tradución non é imposíbel –non o pode ser, pois existe–, mais a distinción entre tradución e adaptación é unha falsidade, ao ser a adaptación o punto de partida inevitábel de todas e calquera tradución.

O grao de adaptación variará, loxicamente, segundo a tipoloxía textual, os sistemas en xogo..., mais sempre existe. Para darmos un exemplo sinxelo, ficou demostrado que é unha práctica habitual e normalizada a substitución de frases feitas específicas do sistema de partida xunto coas súas imaxes por outras equivalentes no sistema de chegada, de as haber. Non é iso máis do que un exemplo claro da tradución como proceso de transferencia intersistémica dinámica.

Ora ben, de aceptarmos que, dun lado, a tradución sempre leva implícita unha adaptación en maior ou menor medida para se axustar ás necesidades inherentes do sistema de chegada e, doutro, que non só se poden senón que se deben axustar as imaxes implícitas en certas expresións feitas etc. para se adecuaren aos equivalentes existentes no sistema meta, cabe preguntarse por que non se procede da mesma maneira cando o tradutor se encontra fronte a unha imaxe xa non virtual senón física que fai parte integral dunha unidade verbo-icónica.

Traducir eficazmente nun mundo onde a imaxe está presente en todos os eidos da vida non pode significar ignorar este facto, senón que o tradutor ten como responsabilidade profesional a de recoñecer o seu valor real para logo procurar efectuar unha transferencia intersistémica efectiva tal e como leva facendo –ou polo menos debería levar tratando de facer– co texto escrito.

Ao fin e ao cabo, a cuestión que se levanta non é outra senón definir cal é a finalidade que se persegue co acto de traducir. Desde a postura que aquí defendemos, esta non pode ser outra do que a procura de tornar transparentes os valores semióticos que revisten os signos formais presentes no texto de partida tal e como se interpretan desde esta estrutura de valores semióticos nun outro polisistema de valores diferentes a través de formas similares ou diferentes segundo for preciso para manter *in tacto* a mensaxe global transmitida a través dos contidos semióticos orixinais. Desde este punto de vista, a ‘adaptación’ débese entender como unha estratexia tradutiva máis e non como unha actividade diferenciada situada alén dos límites do que se entende como tradución, ao ser que sempre existe unha relación cun orixinal alleo á tradución.

Cinxirse e aferrarse ao caduco concepto de ‘fidelidade’ para o orixinal non é máis do que autoimporse unhas chatas desnecesarias que cerran as portas á creatividade precisa para facer do tradutor tecnicamente competente un tradutor excelente canto á súa capacidade de crear textos eficaces e funcionais dentro do polisistema de chegada. É máis, o tradutor ten de asumir que desde o momento en que se muda de idioma, iso

xa de por si leva sempre implícito outras mudanzas consubstanciais que ‘atraizoan’ necesariamente elementos máis ou menos significativos do orixinal. Por exemplo, desde o momento en que se traducen as formas verbais simples e compostas do inglés *he ate* e *he has eaten* pola forma verbal galega simple ‘comeu’, apágase *ipso facto* a distinción implícita nestas dúas formas de partida que, alén da dicotomía simple-composto transmite unha noción de aspecto imperfectivo-perfectivo totalmente ausente da forma única galega que carece de tal distinción gramaticalizada. Trátase dunha perda de ‘fidelidade’ automática e inevitábel, mais non por iso deixa de ser unha falta de fidelidade entendida nos termos tradicionais do termo. En definitiva, achámonos perante un tipo de adaptación inevitábel que se ha repetir de xeito continuo ao longo de calquera tradución sen que iso implique, loxicamente, calquera ‘traizón’ para co orixinal. Se resultan aceptábeis tais adaptacións por seren inevitábeis, habería tamén que reflectir sobre a lexitimidade de efectuar outras mudanzas xa non tan inevitábeis, senón froito da elección consciente do tradutor, porque se encamiñan na mesma dirección e cos mesmos motivos do que o anterior, non se poderían tachar de ‘falta de fidelidade’ ou ‘traizón’ polo simple facto de seren adaptacións.

De facto, o único tipo de tradución realmente ‘fiel’ ao orixinal *stricto sensu*, autenticamente ‘adeuada’ en todo o momento, é a glosa lingüística cuxo propósito consiste, xustamente, en achegar o intérprete alleo ao polisistema de partida ás redes lingüísticas e metalingüísticas que constitúen o polisistema de partida. Neste sentido, fica claro que a tradución propiamente ‘adeuada’ se sitúa no pólo oposto da tradución ‘aceptábel’, pois non pretende ‘domesticar’ texto, senón ‘estranxeirizalo’ ao máximo.

Para darnos un exemplo do devandito, nesta tradución altamente ‘fiel’ dun texto en bretón padrón (texto orixinal tirado de Kervella, 1999: 11) transparéntanse nunha glosa detallada todos os elementos morfolóxicos, léxicos e sintácticos para un lector galego-falante.

1. “Eizhtervezh pe seizh eo emit-hu? a chom c’hoazh da c’hedal.”
2. (Eizh+[dei²+(r)+vezh]) - pe - seizh - (eo) - [(em⁶+it)+hu] - (a+chom) - c’hoazh - (da+/g>c’hedal)
3. /{('oito'+[‘día’+(ligazón eufónica)+(sufixo de duración)]} - ‘ou’ - ‘sete’ - (‘ser’ 3s pres. conx.)} - (‘dicer’+2pl. pres.)+2pl. sufixo pronominal)/ - /(<partícula verbal+raíz verbal ‘ficar’) - ‘ánda’ - (‘a/para’+lenición ‘axexar’/ ‘esperar’)/

Todos/as coincidirían sen dúbida, porén, en cualificaren este tipo de tradución de ‘inapta’ para calquera situación menos a situación moi específica da glosa comparativa lingüística, práctica, por outra parte, de que xeralmente se encargan lingüistas ou filólogos especializados para fins moi determinados e non tradutores.

Fronte á visión, ao noso ver, vetusta no novo século polo que andamos, que concibe a tradución como nun proceso de substitución textual binario onde a misión profesional do tradutor se limitaría a (re)interpretar un texto por outrén, isto é, na procura de interpretar ‘correctamente’ os valores semánticos vinculados polo autor orixinal a través dos seus textos orixinais para transmitirlos coa maior fidelidade posible para un público cuxa lingua dista da que se utilizou para redixir, adaptar implica a plena asunción na práctica do que tantas veces se repite na teoría de concibir o acto de traducir como un proceso triangular onde entran en xogo non só o orixinal e as (supostas) intencións do seu autor e o tradutor, senón un terceiro factor condicionante que desde esta perspectiva resulta fundamental, que é o público receptor, polo que se pode falar non só de traducir por alguén (o autor) senón, e sobre todo, para alguén (o lector):

The function of the source text, and of course also parts of the source text, can or must be changed, depending on the wishes, expectations, needs, etc. of the target readers (Kussmaul 1995: 71).

Poderíase obxectar face a unha tal postura que as eleccións adoptadas polo tradutor residen fundamentalmente nas súas valoracións subxectivas que se imponen sobre as intencións iniciais do autor orixinal que merece o respecto que se manifesta a través da ‘fidelidade’. Cómpre recoñecer, aliás, que tales intuicións sobre as intencións primarias do autor orixinal non son máis do que conxecturas e hipóteses –espérese que fundadas– por parte do tradutor-interpretante.

Face a esta interpretación subxectiva dos posibles motivos primarios do autor orixinal, ao tradutor competente sonlle atribuídos uns coñecementos completos e enciclopédicos a respecto dos valores e estruturas internas dos polisistemas con que traballa. Noutras palabras, ten que realizar o seu traballo dentro dun cadro que se poderá cualificar de competencia ‘biculturalista’, alén da simple competencia lingüística.

Deste xeito, se se elabora con boa fe e partindo de razóns fundamentadas no texto de partida así como nos coñecementos da composición e da estrutura internas dos respectivos polisistemas de traballo, calquera tradución que se realice desde esta óptica será tanto factíbel como lexítima, alén de cuestións de orde máis esotérica sobre a súa ‘corrección’ ou non.

O tradutor só pode ser fiel a si mesmo, isto é, sentirse satisfeito para co seu traballo realizado con profesionalismo e seriedade, e cara ao seu público, pois é este último que lle ha confirmar o éxito do seu texto final ou non, sen que esta valoración resida en ningún momento nunha avaliación do texto traducido respecto do texto orixinal, pois na esmagadora maioría dos casos, o público recep-

tor ten tal condición xustamente polo facto de non poder acceder ao orixinal, tornando por tanto calquera comparanza imposible (ou cando menos moi improbable). O texto traducido hase valorar, ao fin e ao cabo, como un texto autónomo dentro do polisistema literario de chegada e ten que estar elaborado en consecuencia.

Mais, aínda descartando o vello concepto de ‘fidelidade’ cara ao orixinal e asumindo unha nova focaxe orientada cara á integración do texto traducido como un texto funcional e autónomo dentro do polisistema literario de chegada, a diferenza dos textos literarios xerados polos seus propios autores dentro dun polisistema de partida, a tradución distínguese –inclusive defínese– polo facto de ser un texto (re)elaborado unha segunda vez, normalmente por unha persoa diferente do autor orixinal, a partir dun texto orixinal xerado dentro dun outro polisistema de partida. En definitiva, o orixinal existe. Non só existe, senón que serve de base para a reelaboración, resultando indispensable unha exhaustiva análise semiótica dos valores transmitidos polos elementos presentes no orixinal para procurar recrear un efecto no público receptor do texto final similar – ‘equivalente’ – ao producido polo texto orixinal no público orixinal mediante a procura e o establecemento de equivalencias dinámicas.

De todo o anterior decorre o paradoxo característico da tradución que consiste no mito da posibilidade de ficar fiel ao orixinal ao tempo que se muda, pois se non se muda non hai tradución. Ora, tal paradoxo non é máis que superficial e parte do mito socializado que cría a tradución da substitución dun texto para outro igual a pesar de todas as mudanzas efectuadas, quer lingüísticas, quer de referencias culturais...

En suma, a finalidade de calquera tradución sen fins moi específicos é a de conseguir un texto funcional dentro do sistema de chegada partindo dunha relación dinámica de partida preestabelecida polo texto orixinal interpretado desde dentro das relacións polisistémicas correspondentes do sistema de partida e reensablado segundo as necesidades estruturais e semióticas do polisistema de chegada mediante equivalencias dinámicas.

Advogamos, pois, por unha ‘domesticación’ do texto que se ha de entender como a necesidade de o tradutor domar e dominar o texto, sen deixar que o texto orixinal o domine, para se apoderar del, tornándoo seu. O tradutor moderno debe asumir así a responsabilidade e o deber da liberdade creativa que decorre do seu papel non só como interpretante do orixinal senón como reescritor dun texto final que haberá de funcionar autonomamente dentro do polisistema literario de chegada.

Referencias bibliográficas

- Campbell, George, L. (1998): *Concise Compendium of the World's Languages* (Londres: Routledge).
- Chandler, D. (1995): *The Act of Writing. A Media Theory Approach* (Aberystwyth: Universidade do País de Gales).
- Kervella, Goulc'han (1999): *E vrenigenn diwezhañ* (Brest: Al Liamm–An Here).
- Kim-lung Au, Kenneth (1999): “Cultural Transfer in Advertisement Translation”, en Caillé, P.-F. (fond.): *Babel. Revue Internationale de la Traduction*, 45: 2 (97-106) (Amsterdam/Filadelfia: John Benjamins).
- Kussmaul, P. (1995): *Training the Translator* (Amsterdam/Nova Iorque: John Benjamins)
- Moravia, A. / Mancuso, G. [trad. Palao, J.] (1975): *Mao Tsé-Tung* (Madrid: Júcar).
- Pinker, S. (1994): *The Language Instinct* (Londres: Penguin).
- Rabadán, R. (1991): *Equivalencia y traducción. Problemática de la equivalencia translémica inglés-español* (León: Universidad de León).
- Sampson, G. (1985): *Writing Systems* (Londres: Hutchinson).

César Candelas Colodrón, *O Cronicón de Hidacio, bispo de Chaves, Noia, Toxosoutos, 2004 (Serie Trivium 13).*

José Antonio López Silva, *A Crónica de Idacio da Limia, bispo de Chaves, Deputación Provincial de Ourense, 2004.*

Á curiosa circunstancia de que a *Crónica* de Hidacio, de interese xeográfico moi localizado, coñecese ata o momento tradución a varias linguas europeas, mais non ao galego, vén de sumarse agora a feliz casualidade da publicación simultánea de dous traballos que corríxen esa deficiencia. Os dous comparten algunhas características en común –abordan en senllos estudos introdutorios os aspectos fundamentais sobre o autor e a súa obra e as traducións están feitas sobre a edición máis recente do texto latino, a de Burgess–, mais son complementarios.

O traballo de C. Candelas, de corte basicamente histórico, presenta a tradución da *Crónica*, precedida dun estudo preliminar en que, tras unha Introducción (pp. 9-13) adicada ás publicacións máis significativas dos últimos decenios sobre a materia, aborda os aspectos biográficos e a actividade política e relixiosa de Hidacio en dous apartados, “O xove Hidacio” e “Hidacio bispo”. Canto á biografía, os datos que achega son os xa coñecidos, e extraídos da propia *Crónica*, pero ofrece as interpretacións máis recentes e, sobre todo, chama a atención sobre os prexuízos ideolóxicos de certos traballos, como os de C. Torres.

A parte mais novidosa é a que se refire á etapa do episcopado do autor, con epígrafes adicadas ao contexto xeográfico-histórico, ao papel de Chaves como centro de poder independente, a Hidacio como líder da comunidade e á ideoloxía hidaciana. Neste último apartado insiste Candelas no perfil pouco “belixerante” que dos bárbaros dá Hidacio e nos seus posibles vínculos coa familia de Teodosio. Bótase en falta neste lugar unha mención ou unha valoración da caracterización que algúns traballos fixeron do autor como protonacionalista galaico, coa que Candelas parece mostrarse de acordo na Introducción (p. 11).

O traballo de López Silva é de mais ampla aspiración e combina a orientación histórica cunha rigorosa análise filolóxica, sen fuxir da toma de partido a propósito dos aspectos mais discutidos da obra. A introdución está dividida en dous apartados, adicados ao contexto histórico e biográfico de Hidacio e á propia *Crónica*, respectivamente. No aspecto biográfico proporciona os mesmos datos que a publica-

ción precedente, ocupándose de cuestións coma o lugar de nacemento do autor, a súa viaxe a Oriente e vida política. O estudo da *Crónica* está focado desde o punto de vista do xénero literario a que pertence, cunha análise pormenorizada da tradición que arranca de Eusebio e Xerome. O tradutor coñece moi ben os traballos existentes sobre a materia e fai unha valoración crítica das súas aportacións, sobre todo en aspectos sometidos a discusión, coma o problema do universalismo e o localismo. O punto de vista que adopta o tradutor nalgunhas cuestións en que non é doado manter a equidistancia entre extremos, é sempre ponderado, lonxe de sentimentalismos ou de divulgacións superficiais.

A tradución vai acompañada dun amplísimo comentario histórico, a que seguen o texto latino de Burgess e mais unha recompilación bibliográfica, tan exhaustiva como excesiva, pois, ao lado de obras de consulta imprescindible, contén referencias moi marxinais ou apenas relacionadas co tema obxecto de estudo. Cerra o volume unha reprodución facsimilar da tradución ao castelán feita por Marcelo Macías e publicada no *Boletín de la Comisión Provincial de Monumentos de Orense*. Son dous acertos o feito de incluír o texto latino, aínda que sería de consulta mais cómoda se estivese presentado en páxinas enfrontadas coa tradución, e o de reproducir o traballo de Macías, de difícil acceso, mesmo para o especialista. En resumo, un libro completísimo, que vai ser de cita obrigatoria para quen volva sobre a *Crónica* hidaciana.

As dúas traducións son correctas e mesmo elegantes, se temos en conta as limitacións do propio texto latino, incardinado nun xénero, en xeral, aliterario. Neste sentido, o prefacio da *Crónica*, de carácter moi retórico, con períodos difíciles e mesmo ás veces susceptible de ser interpretado en sentidos diversos, contrasta notablemente co carácter esquemático da parte propiamente histórica. E é nel onde se poñen de manifesto os problemas de calquera tradutor. Seleccionarei algúns exemplos significativos das traducións de Candelas e López Silva, comezando polo inicio do citado prefacio: *HVCVSQVE A SANCTO HIERONIMO ET IPSO, SICVT IN CAPITE ISTIVS uoluminis praefatio prima declarat, cognomine Eusebio historia in aliquantis Hispaniarum prouinciis conscripta retinetur; cui siquid postea subididit in locis quibus deguit, certo stili studio declaratur*. Hidacio, como farán despois outros continuadores do mesmo corpus cronístico, deixa constancia explícita de que ata ese punto chegaba a obra escrita por Xerome, que recibiu tamén o sobrenome de 'Eusebio', *sicut in capite istius voluminis praefatio prima declarat*. Hai que pór en relación esta frase co verdadeiro carácter do texto hidaciano, en que tanto insistiu P. Nautin e acertadamente enfatizado por López Silva, o de ser unha continuación da *Crónica* de Eusebio-Xerome. A única interpretación posible é que *in capite istius uoluminis* se refire ao volume que contiña a *Crónica* de Xerome, que no prefacio indicaba o nome e o sobrenome de seu autor. Así intenta reflectilo a tradución de C. Candelas, "Ata aquí a historia conservada nalgunhas provincias

de Hispania e escrita por San Xerome, tamén coñecido, tal e como se sinala na introducción ó inicio deste libro polo alcume de Eusebio”. A tradución de López Silva, en cambio, resulta máis problemática, “Ata aquí está escrita a historia do santo Xerome, de sobrenome Eusebio, conservada nalgunhas partes de Hispania, segundo declara o primeiro prefacio deste libro”. Tal formulación deforma o sentido do texto ata facelo dificilmente comprensible, ao romper a asociación entre o que se afirma (‘de sobrenome Eusebio’) e o lugar onde se afirma (‘segundo declara ... libro’). Unha simple variación de orde de palabras bastaría para restablecer o verdadeiro sentido do texto: “Ata aquí está escrita a historia, conservada nalgunhas provincias de Hispania, do santo Xerome, de sobrenome Eusebio, segundo declara o primeiro prefacio deste libro”.

A frase seguinte do mesmo prefacio tampouco está exenta de problemas, ata o punto de que dá a impresión de que ningún dos dous tradutores se sentiu cómodo perante a enigmática frase *siquid ... declaratur*: “se lle engadiu algo despois nos sitios onde viviu, ponse de manifesto nun preciso estudio da súa obra” (Candelas) / “Se nos lugares onde viviu lle engadiu algo máis, certamente iso quedaría declarado pola súa obra literaria” (López Silva). O parágrafo presenta problemas interpretativos. Hidacio deixa no aire a dúbida sobre a existencia dalgunha versión da obra de Xerome máis completa que a que el continúa, é dicir, presenta a posibilidade de que o autor continuase a súa tarefa nos lugares en que viviu. Que quere dicir *certo stili studio declaratur*? Ao meu xuízo, estamos diante dunha pregunta retórica, para a que se desprende do texto unha resposta afirmativa: “se despois lle engadiu algo ... dedúcese da súa afición certa pola escritura”. O inicio do parágrafo 4 reforza esta interpretación, *Esto ut in sanctis quibus deguit Hierosolimorum locis ... forte quam plurima de his que sunt insecuta subiecerit, quia aud umquam, dum ualuit, a diuerso stili opere cessauit* (tradución na p. 67 Candelas, onde hai un erro de numeración do parágrafo, e na p. 68 López Silva).

Na continuación do mesmo parágrafo 4, a tradución de López Silva presenta certas dificultades. O texto latino di o seguinte: *Si tum proprio operi quod subdidit aliqua subiunxerit, apud eos, ad quos scriptorum eius omne opus uel summa peruenit, certa et plena cognitio est*. Hidacio afirma claramente que só aqueles aos que chegou a obra completa de Xerome ou unha suma dela teñen certeza plena sobre a existencia ou non de adicións ulteriores do autor. López Silva parece non decatarse disto e transforma unha afirmación rotunda en dúbida (ou en interrogación retórica con resposta negativa?): “Se entón tivera engadido algo máis a esta continuación da súa propia obra, existiría o coñecemento certo e pleno disto entre aqueles ós que lles chegou todo o conxunto dos seus escritos ou unha gran parte deles”.

Canto á tradución, á marxe destes exemplos illados, que atopan fácil explicación polas propias características dun texto que admite interpretacións diversas, o traballo de C. Candelas presenta un reparo serio na tradución das datas do mes. Entre

manter o sistema de datación latino, facendo referencia ás calendas, nonas ou idus –como fai, por exemplo, López Silva– ou utilizar o sistema de equivalencias no calendario actual, é preferible o segundo, dado que, para un potencial lector non informado sobre o modo de datar en latín, facer referencia ás calendas, nonas ou idus non lle permite entender o texto. Candelas trata de evitar o problema dando as dúas (a do calendario actual entre paréntese), mais a forma de se referir á datación propiamente latina é aberrante. Datas como “na 13ª calenda de Agosto (19 de Xullo)” (p. 75), “na 14ª calenda de Novembro (19 de Outubro)” (p. 81), “na 5ª calenda de Xuño (25 de Maio)” (p. 90) ou “na 6ª nona de marzo (2 de Marzo)” (p. 92) son só algunhas mostras do que non debe facerse. A “13ª calenda de Agosto” é un día do mes de Agosto ou de Xullo? Un lector medio tería moitas dificultades para sabelo. En lugar desas referencias tan confusas (e erróneas, porque as calendas eran o día primeiro de cada mes, e non había varias ao longo del), tería que dicir “o día 13 antes das calendas de Agosto (19 de Xullo)”, “o día 14 antes das calendas de Novembro (19 de Outubro)”, “o día 5 antes das calendas de Xuño (25 de Maio)” ou “o día 6 antes das nonas de Marzo (2 de Marzo)”, ou ben dar unicamente a referencia actual. O lector ben merece que non se lle compliquen as cousas innecesariamente.

O libro de López Silva, pola súa parte, presenta, desde o punto de vista formal, algún lapso no galego, fácilmente explicable polo estado aínda magmático da lingua, e ofrece na introdución os termos gregos sen os signos gráficos de espíritos e acentos, deficiencia que se tería evitado recorrendo a conxuntos de caracteres que conteñen tales signos.

Á marxe das suxestións, que non críticas, ata aquí citadas, cabe felicitar aos tradutores por dous magníficos traballos que actualizan un texto por veces problemático, pero fundamental para a historia da Gallaecia da Antigüidade Tardía.

Arsenio Ferraces

Goretti Sanmartín Rei, *Lendo nas marxex. Lingua e compromiso nos paratextos (1863-1936)*, A Coruña, Espiral Maior, 2002.

Felizmente, son xa varios e importantes os traballos realizados desde unha perspectiva sociolingüística que centran a súa análise na situación do galego durante os séculos XIX ou XX e que a basean no estudo de diferentes tipos de textos, a incluíren tamén aqueles propiamente literarios que, dunha forma ou doutra, conteñen información sociolingüística. Aínda que nalgún caso se acudía puntualmente a algún prólogo, nomeadamente os rosalianos, porén, non se lle tiña prestado á instancia prefacial unha atención central no evoluir da situación sociolingüística en que a produción escrita en galego se desenvolveu durante os dous séculos precedentes. A profesora Goretti Sanmartín Rei veu a encher ese oco no ano 2002 coa publicación de dúas importantes monografías derivadas da súa ben documentada tese de doutoramento lida na Universidade da Coruña, realizando desa forma un contributo fundamental aos estudos de sociolingüística galega.

Estes traballos son *Lendo nas marxex. Lingua e compromiso nos paratextos (1863-1936)*, obra en que nos centraremos, e *Os (pre)textos galegos (1863-1936)*, perfecto complemento do anterior que, alén dunha extensa e rigorosa introdución arredor dos prólogos (concepto e importancia, relación co ensaio, características e etapas), escolma nin máis nin menos que 95 dos escritos só en galego, deixando fóra, por evidentes razóns de espazo, outros paratextos na nosa lingua de indubitábel interese, así como todos aqueles redixidos noutras linguas e que tamén teñen relevancia para o coñecemento do contexto sociolingüístico en que os textos se desenvolveron; un dos grandes contributos deste traballo é achegarnos textos até agora xeralmente descoñecidos e de difícil localización, sobre todo aqueles de autoría menos coñecida. Con estas dúas obras, que son merecentes do cualificativo de novidosas e orixinais, pódese afirmar que a sociolingüística galega dá un paso adiante importante; inclusive noutras linguas con máis tradición investigadora e máis posibilidades non se teñen producido na mesma medida estudos destas características ou similares.

Centrándonos, pois, en *Lendo nas marxex*, libro que xa no subtítulo aparece cronoloxicamente delimitado ao período que vai de 1863 a 1936, anos ben significativos cuxa relevancia non é preciso comentarmos, e tematicamente dirixido ao estudo do compromiso lingüístico contido nos paratextos, convén comezarmos por afirmar, como nos di a autora no seu propio “limiar”, que estas unidades textuais contribuíron de modo decisivo ao proceso de normalización lingüística, tanto divul-

gando as ideas nacionalistas, como reflectindo unha mudanza na perspectiva de confinar o galego ao ámbito rural e a un discurso diglósico, a superaren así mesmo “o vello prexuízo que establecía un discurso en prosa en español antecedendo un libro de poemas ou unha peza dramática en galego e dignificando e prestixiando o desenvolvemento dunha literatura cuxo reto maior era estar dirixida a un público lector que ora non sabía ler, ora tiña o español como lingua de cultura” (p. 9). A función principal dos paratextos estudados nesta obra, nun período de acerbadísimo conflito diglósico, era defenderen e arrouparen os propios textos, procurando anticipárense ás posibles críticas ou xustificando a razón de estaren escritos en lingua galega. Convertíanse así os paratextos, a comezar polo título e continuando polas dedicatorias ou os prólogos alógrafos ou de autor para terminar nas notas e epílogos, en escudos protectores dos textos nun contexto hostil ao emprego do galego, onde a aparición dun libro nesta lingua aínda supuña unha grande anomalía que era preciso, dalgunha forma, xustificar.

A obra estrutúrase en catro capítulos, a constituíren os dous primeiros o verdadeiro corpo do traballo e os dous últimos a parte complementaria que fornece elementos imprescindibles de lectura, información e consulta. Así, o capítulo primeiro (“Os paratextos. Definición e tipos”, pp. 11-76), após unha introdución en que a autora, a se apoiar en Genette, de quen toma a propia denominación de ‘paratextos’, procura unha definición e caracterización deles, distinguindo entre ‘peritextos’ (aparecen no volume: título, subtítulo, intertítulo, prólogo etc.) e ‘epitextos’ (están fóra do texto: anuncios en revistas ou xornais, entrevistas co autor ou autora, críticas da época, correspondencia privada etc.), pasa a tratar sobre cada un dos constituíntes dos paratextos: autores e autoras, títulos, dedicatorias, prólogos, advertencias editoriais, epílogos, notas, outros elementos paratextuais de carácter verbal e elementos paratextuais non verbais. En cada un destes apartados aprofúndase na súa caracterización tipolóxica e destácase a información sociolingüística que conteñen, con frecuente exemplificación de diferente autoría e época dentro do período en que a obra se enmarca, levantando interesantes citas de escritores/as e obras que pasaran anteriormente desapercibidas, para alén doutras xa clásicas na nosa tradición, que loxicamente non se podían pasar por alto.

O capítulo segundo (“A situación do galego na época”, pp. 77-218) constitúe, tanto polo lugar que ocupa na obra como pola súa extensión e contido, o núcleo central do traballo, pois nel procurábase un estudo da situación lingüística durante o período predeterminado con base nos datos tirados da instancia prefacial e paratextual en xeral, pondo en confronto as informacións obtidas por esta vía con aquelas procedentes doutros estudos, manexando neste sentido a autora adecuadamente a bibliografía existente ao respecto. Após unha longa introdución á situación do galego na época (pp. 77-91), dedícanse apartados á cuestión do número de falantes (pp. 91-111), ao confronto entre a aldea e a cidade canto ao uso do galego (pp. 112-128),

á utilización da lingua galega noutros ámbitos de uso, tanto orais (conferencias, uso privado etc.) como escritos (pp. 129-135), ao estado da literatura galega (pp. 135-148) e a consideracións sobre o modelo de lingua utilizado polos escritores, apartado que presta atención á conformación dun modelo de lingua literaria dunha forma demorada e ben documentada (pp. 148-200) e ao sempre polémico achegamento a Portugal (pp. 200-218). Pásase, pois, neste capítulo central da obra revista ás cuestións lingüísticas máis actuais e tamén problemáticas na altura, algunhas das cales aínda o continúan a ser na actualidade, pois por veces semella que temos avanzado pouco ou nada en certos aspectos da realidade lingüística. Noutros, felizmente, déronse pasos importantes, a pesar de que moitas das grandes cuestións que afectan o futuro do noso idioma fican aínda pendentes dunha solución satisfactoria. Este capítulo fornece informacións e reflexións dignas de seren tomadas en consideración para unha adecuada orientación dos problemas que subsisten e que dificultan a normalización social do galego.

No capítulo terceiro tíranse por extenso (pp. 219-234) as conclusións do traballo; mais, no canto de optar por unha concisa enumeración destas, como adoita ser habitual, a autora desenvolve un capítulo resumativo en que vai tecendo o evoluir continuo do discurso paratextual en paralelo co propio desenvolvemento do movemento galeguista, resaltando a importancia da aparición do nacionalismo como corrente política necesitada de modernizar o discurso para o facer atractivo á sociedade e que, ao mesmo tempo, lonxe de o separar do contorno ideolóxico da época, o utilizou como instrumento de propagación do pensamento galeguista. Os elementos paratextuais caracterizáronse, así, polo “seu didactismo, o afán proselitista e o carácter pedagóxico, nun intento de suplir todo aquilo que a cultura oficial lle negaba ao noso idioma, producindo un efecto amplificador, cualitativa e cuantitativamente, dos usos da lingua e facilitando unha visibilidade social até o momento inexistente” (p. 224). Os paratextos tiveron de se atribuír, deste modo, funcións que nunha situación de normalidade lles corresponderían a outras instancias, adquirindo por iso unha relevancia especial e a ocuparen as ideas sobre a lingua un lugar de preeminencia neles, feito que de seguro non tería acontecido nunha literatura normalizada. Sintetiza tamén a autora neste capítulo a evolución dos títulos (onde se acabará por xeneralizar a forma *galego* desde o predominio inicial de *gallego*), dos subtítulos ou da instancia prologal, elemento básico no diálogo entre emisor e receptor que adquire unha grande importancia, pois “a través dos prólogos temos coñecemento das polémicas lingüísticas, literarias e políticas, e, en situación de conflito lingüístico, estes supoñen un dos recursos que o autor ou edior ten para difundir a necesaria (auto)xustificación, servindo de escudo protector dos moi diferentes ataques que sen dúbida o libro vai provocar” (p. 232). Por outra parte, o carácter preceptivo na procura da fixación dun modelo de lingua literaria digno, que supuxese tamén a necesaria unificación ortográfica e morfolóxica, foise facendo moi presente nos paratextos do último período estudado pola autora.

O capítulo cuarto dedícase ás referencias bibliográficas (pp. 235-257), distinguindo por un lado os ‘estudos’, onde se recollen todos os traballos citados no desenvolvemento da obra, tanto referidos ao ámbito exclusivamente galego como aqueles de carácter xeral que son utilizados como referentes teóricos na análise paratextual, e por outro lado os ‘textos’, onde a autora nos fornece unha ampla listaxe (pp. 243-257), alfabeticamente ordenada segundo a autoría, dos paratextos aparecidos nas obras do período 1863-1936, de grande utilidade por nos descubrir textos e mesmo obras pouco ou nada coñecidas que, se ben non destacaron precisamente pola calidade literaria, e por iso mesmo foron xeralmente ignoradas, si posúen grande interese para a análise sociolingüística que na obra se pretende e inclusive tamén para a análise estritamente lingüística no caso dos textos en galego, pois neste caso recóllense así mesmo os paratextos noutras linguas. As propias características da evolución do noso sistema lingüístico-literario non nos deben permitir obviarmos estas fontes de información e análise, en ocasións imprescindíbeis, perante a ausencia doutros medios existentes en contextos de normalidade cultural. Así pois, o obrigado capítulo bibliográfico adquire nesta obra un valor especial por fornecer aos estudiosos e investigadores unha información sistemática e novidosa, dificilmente achábel nos repertorios bibliográficos preexistentes.

Após esta breve síntese do contido da obra, e véndoa no seu conxunto, podemos cualificala como oportuna, necesaria e orixinal, pois converte as unidades paratextuais, xeralmente preteridas na análise sociolingüística, en elementos centrais de estudo sobre a situación do galego nese período histórico, resgatando fontes de información até o momento esquecidas. Nun contexto como o galego os paratextos desempeñan unha función especial, tamén constatábel noutras literaturas periféricas en que o seu desenvolvemento se ve fortemente condicionado pola utilización diglósica do idioma, converténdose en principais receptáculos de ideoloxía e mesmo a condicionaren a lectura dos textos literarios, ao se interporen no circuíto comunicativo existente entre locutor ou autor/a e alocutario ou lector/a. En certo modo, os paratextos veñen a furtar as múltiples posibilidades de expresión que normalmente posúe un texto, que desta maneira “ve limitadas as súas lecturas nun contexto diglósico” (p. 225), fenómeno ben analizado por Antón Figueroa en *Diglosia e texto*. Resulta evidente que se un texto é lido en función do código lingüístico utilizado ou do seu valor simbólico-patriótico, como con frecuencia se suxire nos paratextos, iso está a interferir o seu normal funcionamento como produto estético. Coñecermos esas circunstancias que rodearon a produción textual axudaranos a entender mellor a situación por que pasou a nosa lingua e a evolución do noso sistema literario e mesma da conciencia social no seu conxunto. Ou, como con máis fortuna afirma a autora, é “preciso detérmonos neses outros textos que tamén constrúen o noso discurso e, a través deles, percorrermos as beiras do río nun camiño cara a atrás que ten a intencionalidade manifesta de se erixir nun compromiso co futuro do noso idioma” (p. 10).

Así pois, este traballo de Goretti Sanmartín Rei, complementado co outro volume arriba citado da mesma autoría, vén a deitar unha nova luz sobre a nosa historia lingüística e literaria, cumprindo perfectamente o obxectivo, fixado pola autora no seu propio “pretexto” ou “limiar”, de realizar unha lectura sociolingüística dos principais elementos paratextuais do período estudado, mediante a cal se demostrase que os paratextos forman parte fundamental da historia da lingua e da literatura da Galiza e se constatase o seu especial funcionamento nunha situación de conflito diglósico como a que aquí se deu e aínda se está a dar. Obxectivos, pois, cumpridos de forma clara nesta obra, segundo o lector ou lectora poderá comprobar se optar pola súa lectura, especialmente recomendábel pola novidade e orixinalidade que supón no ámbito dos traballos sociolingüísticos e particularmente gratificante pola súa excelente redacción, cun modelo de lingua que vén a demostrar que, dentro das marxes da norma, é posíbel a utilización dun galego depurado e auténtico que non se afaste no substancial das súas propias raíces e do sistema en que está inserido.

Xosé Ramón Freixeiro Mato

Xosé Ramón Freixeiro Mato, *Gramática da lingua galega. IV. Gramática do texto*, 2003, Vigo, A Nosa Terra.

Despois da monumental obra en tres volumes publicada por A Nosa Terra e intitulada *Gramática da lingua galega* (I, *Fonética e Fonoloxía*, 1998; II, *Morfosintaxe*, 2000; e III, *Semántica*, 1999), e após tamén unha síntese deses tres tomos destinada a un público máis amplo e dada a lume pola mesma editora (*Manual de gramática galega*, 2001), o profesor Xosé Ramón Freixeiro Mato volta de novo á actualidade editorial do país por ser o autor da primeira teorización feita para o galego dentro do marco teórico coñecido como Lingüística do Texto: *Gramática da lingua galega. IV. Gramática do texto*. Este contributo, que saíu á rúa, á semellanza dos anteriores, das mans d'A Nosa Terra no ano 2003, foi entendido como unha outra parte do proxecto orixinal en tres volumes, xa que, como o propio autor sinala, “a gramática do texto non vén a ser máis do que outra parte da gramática dunha lingua, aquela que toma o texto como unidade de análise e estudo, superando o marco da cláusula en que tradicionalmente a gramática se detiña” (p. 11).

A nova contribución de Freixeiro Mato á lingüística galega aparece estruturada en nove temas que constitúen, por súa vez, os nove centros de interese que desenvolve o investigador da Universidade da Coruña. No primeiro deles, que cumpre a función dun necesario preámbulo ás oito unidades restantes (“A lingüística textual”, pp. 13-26), Freixeiro desenvolve un percurso histórico polo desenvolvemento da lingüística do texto como máis unha posibilidade de análise da lingua, salientando os antecedentes, as fases na evolución desta disciplina etc. “Discurso e texto” (pp. 27-32), que corresponde ao segundo dos temas, céntrase na esguilante (in)distinción entre ambos os conceptos partindo da equivalencia discurso-oralidade / texto-escrituralidade, que o profesor Freixeiro resolve a indicar que a tendencia dominante na actualidade é a “denominación única de texto, escrito ou falado” (p. 30). Co título “A diversidade dos xéneros discursivos” desenvólvese a unidade seguinte (pp. 33-41), en que se pasa atenta revista ás diferentes propostas existentes ao respecto, que presentan entre si importantes diferenzas, mais que, en todo o caso, e como sinala o autor, “cando unha persoa escolle un texto está a escoller unha maneira de actuar sobre alguén” (p. 39), onde se pondera o efecto que puiden ter un determinado texto sobre o eventual público ouvinte ou lector. O cuarto tema está consagrado aos tipos de textos existentes (pp. 44-62), en que, se ben é certo que non existen tipos textuais puros, se ofrece unha clasificación en textos argumentativos, explicativos, descritivos, narrativos e conversacionais. As relacións entre

contexto e texto, doutro lado, son analizadas na unidade seguinte (“Texto e contexto”, pp. 63-76), onde se tamén aborda a situación particular da Galiza como consecuencia da diglosia existente. “A organización da información” (pp. 77-102) é o tema seguinte, que se centra no estudo dos recursos lingüísticos que unha persoa utiliza para tonsorar o seu particular punto de vista na organización do texto. Doutra parte, o tema sete estuda a cohesión (pp. 103-135), o oitavo analiza nos marcadores discursivos (pp. 137-174) e o derradeiro destínase á análise da coherencia (pp. 175-185). A obra finda coas referencias bibliográficas (pp. 187-194) e cuns útiles índices (pp. 195-208), un de materias e outro de autores.

Desta proposta de estruturación temática interesa resaltarmos algúns aspectos, como poden ser, a modo de exemplo, a decisión de tratar por separado a coherencia e a cohesión. Non resulta fácil determinar os límites precisos de ambos os conceptos, e boa proba diso é que, como indica Freixeiro Mato, “hai autores que integran a coherencia na cohesión e outros que inclúen a cohesión na coherencia” (p. 179), sen faltaren investigadores que as tratan por separado, como Beaugrande / Dressler (1997). Mais, en harmonía con estes estudiosos ou con Koch / Travaglia (2002), o autor decide atinadamente distinguilas, xa que, como se indica neste traballo, a cohesión é “explicitamente revelada através das marcas lingüísticas”, que son “índices formais na estrutura da seqüencia lingüística e superficial do texto”. Congruentemente, pois, con esta decisión, desenvólvese os tipos de coherencia e cohesión, salientándonos nesta última, entre outros, os recursos gramaticais e lexicais; esta perspectiva incide nas responsabilidades textuais dos pronomes como elementos que serven para recobren información xa indicada (déixe textual ou anáfora), e permite entendérmolos como unidades fundamentalmente operativas no texto, superando así visións tradicionais, hoxe felizmente cada vez menos recorrentes, que os identificaban como as “palabras que substitúen os nomes” (*Os alumnos_x coñeceron uns turistas moi agradábeis_z. Estes_z pretendían convidalos_{xi} a xantar*). Aliás, as relacións semánticas de hiperonimia, hiponimia e cohiponimia, dun lado, e sinonimia, doutro, analizábeis dentro dos recursos lexicais, permiten, así mesmo, comprendermos esas nocións semánticas nunha dimensión máis ampla, pois serven, outravolta, como mecanismos que dotan o texto de cohesión (*Vin un paxaro_x onte e era unha ave_{xi} moi bonita*).

Máis un importantísimo acerto deste traballo, que é común cos tres volumes anteriores, é o feito de o autor estar a dialogar permanentemente coa literatura lingüística existente sobre os diferentes temas que desenvolve. Deste modo, polas páxinas desta *Gramática do texto* podemos compartillar e confrontar opinións con Escandell Vidal, Conte, Villaça Koch, Maingueneau, Marcuschi, Schiffirin, Sperber / Wilson, Van Dijk etc., que acompañan o profesor Freixeiro Mato ao longo das duascenas páxinas que conforman este tomo cuarto. Convén facermos notar que a tradición gramatical galega, quitando algún caso excepcional e nomeadamente até

a publicación dos tres volumes anteriores da obra de Freixeiro Mato, obviou ese constante cotexamento, esa continua imbricación reflexiva cos principais autores das materias ou aspectos que se estudan.

Outrosí, ao falarmos de bibliografía e de principios comúns co proxecto global de gramática galega en que se insire este traballo, non podemos deixar de comentar máis unha feliz concomitancia, cal é a utilización das gramáticas galegas existentes, desde as primeiras, as do século XIX, até ás últimas novidades editoriais, xa nos comezos do XXI. Hai abundantes exemplos en que se pode verificar ao longo da obra estoutro diálogo coa nosa tradición filolóxica, que, a pesar das súas eivas e características, é empregada con total acerto. Achamos unha boa proba disto no tema dedicado aos marcadores discursivos, en que figura un apartado consagrado a ver como os elementos que poden desenvolver estas funcións textuais (tradicionalmente coñecidos como “partículas” ou “palabras invariábeis”, como as preposicións e conxuncións, certos adverbios, interxeccións etc.) foron entendidos por autores como Saco Arce (1868), Lugrís Freire (1931), Carré Álvarellos (1967), Carballo Calero (1979), Álvarez / Regueira / Monteagudo (1986), Costa Casas / González Refoxo / Rábade Castiñeira / Morán Fraga (1988) ou Álvarez / Xove (2002). Pola lectura dese apartado podemos descubrir o modo en que determinadas palabras, que non encaixa(ba)n satisfactoriamente nas clasificacións tradicionais, lles supoñen aos autores respectivos certos problemas de identificación, precisamente pola súa singular natureza: Freixeiro Mato (p. 153), por exemplo, cita o caso de Saco Arce (1868: 134), quen se refire a formas cal *madia*, *ande* e *nô* como elementos “de natureza ambigua, participando á la vez de adverbios, conxunciones e interjecciones, siendo verdaderas oraciones elípticas”.

Está tamén presente nas páxinas deste tomo cuarto unha firme vontade de rexeneración e de dignificación do propio corpus lingüístico do galego, que se traduce, do mesmo modo que nas restantes partes da obra, nunha priorización das estruturas morfosintácticas máis idiosincráticas e xenuínas do galego, tales como o futuro de subxuntivo, o infinitivo flexionado, o infinitivo xerundial, a próclise do pronome átono rexido por preposición etc., isto é, empréganse propositadamente aqueles trazos que serven para identificaren o sistema lingüístico galego-portugués a que pertence o noso idioma. Así, a lectura dos traballos de Freixeiro Mato, alén de se constituír cientificamente nun material de primeira orde para coñecermos a realidade da nosa lingua, permite profundarmos nas construcións e nos xiros e máis autóctones do galego, sempre os empregando de acordo coa nosa tradición lingüística milenaria.

E esa constante preocupación por un modelo de lingua culta, apto para todo o tipo de contextos, percíbese tamén no reprobamento que Freixeiro fai contra certa caste de marcadores discursivos existentes no galego como produto da interferencia do español (pp. 162-166), que corresponden, de acordo co profesor coruñés, a catro

tipos: a) aqueles tirados directamente desa lingua (**bueno*, **o sea* etc.); b) os parcialmente adaptados ao galego (**é decir*, **de todos xeitos* etc.); c) os de dubidosa galegitude (*?sen embargo*, *?en cambio*); d) aqueles en que se detecta maioritariamente a forma coincidente coa española (*en realidade*, *con respecto a* etc., en vez de *na realidade*, *a respecto de* etc.); e) os de escaso uso no galego moderno (*porén*, *en canto* etc.); e f) aqueles que se comportan como sinais conversacionais calcados do español (**veña*, **vale* etc.). Convén notarmos a pertinencia destas indicacións, pois no discurso oral galego, para alén doutros tipos de españolización rexistrábeis (lexical, morfolóxica, sintáctica etc.), tamén se encontran esas unidades; mesmo algunha delas, como **bueno* ou a súa variante popularizada e vulgarizada **boeno*, gramaticalización dun adxectivo castelán, aparecen constantemente na oralidade dos nosos días, mesmo na daquelas persoas con consciencia lingüística. A polifuncionalidade de tales marcas discursivas, que as fai apareceren con significacións textuais diferentes nunhas proporcións de uso certamente altas, concédelles a eses marcadores españolizados, nas palabras de Freixeiro Mato, “unha especial relevancia no avanzado proceso de castelanización en que a lingua galega se acha na actualidade e a que os converte en símbolos representativos da desnaturalización que está padecer o galego” (pp. 164-165).

Até o de agora estivemos a comentar algúns puntos en común aos catro tomos da obra, que son, entre outros factores, os que lle conceden a coherencia necesaria a pesar de ser un proxecto cuxas partes saíron do prelo en anos distintos. Porén, existen algunhas diferenzas con relación ás entregas anteriores, o que sen dúbida é explicábel pola propia xestación e conformación deste cuarto volume. Deste modo, o público lector que coñecer os outros tres volumes notará en falta, quizais, a recorrencia aos textos literarios que se encontra na *Fonética e Fonoloxía*, na *Morfosintaxe* e na *Semántica*, o que ten a súa explicación, como declara o propio autor na “Nota introdutoria”, na inicial concepción deste tomo, que en principio non foi pensado como unha continuación das tres partes anteriores. Así mesmo, congratulándose o autor polas últimas modificacións do estándar galego, aplica congruentemente os seus acordos e indicacións, e é por iso que entre os tres tomos anteriores e este se deixan ver algunhas mudanzas fundamentalmente morfolóxicas, o que non resta consistencia, no entanto, á procura dun modelo de lingua dignamente inserido no espazo lingüístico galego-portugués.

Outro dos asuntos a que non queríamos deixar de facer mención é a relación entre texto e contexto, e en especial ás circunstancias sociais, culturais, lingüísticas e políticas que determinan a produción de textos na Galiza, mesmo os literarios (“O contexto sociolingüístico. Diglosia e texto”, pp. 70-76). Nunha situación como na que se encontra galego, os vínculos evidentes entre autor/a e lector/a, ou, se se preferir, entre locutor/a e ilocutor/a, adquiren novos significados e valores e poden mesmo hipertrofiar a creación e a recepción do texto, secundarizar a vontade

estética, abretemar o significado real do texto etc. Ese mesmo contexto histórico e sociolingüístico é o que explica, sen irmos máis lonxe, a demora no aparecemento das nosas primeiras gramáticas, e, indirectamente, é tamén o motivo de por que até o ano 2003 non contou o galego cunha ferramenta como esta *Gramática do texto* que estamos a comentar.

Xa para finalizarmos estas liñas, e a modo de conclusión, non podemos facer cousa que non sexa dámoslle os parabéns ao autor por, máis unha vez, situar o galego no mesmo nivel do que outras linguas do seu contorno no que se refire á reflexión lingüística desde un ámbito innovador de análise gramatical, máxime, aínda, a termos en conta que a nosa lingua non contaba con ningún traballo de semellantes características e que, grazas a este novo contributo, se cobre un relevante baleiro dentro dos estudos de lingüística galega. Así pois, Xosé Ramón Freixeiro Mato convértese con este traballo no precursor da investigación lingüística aplicada ao galego desde a disciplina da Gramática do Texto, como xa antes o foi por publicar os tres tomos da máis completa, documentada e rigorosa *Gramática* das existentes en galego.

Xosé Manuel Sánchez Rei
sanrei@udc.es

Referencias bibliográficas

- Álvarez, R. / Regueira, X. L. / Monteagudo, H. (1986): *Gramática galega* (Vigo: Galaxia).
- Álvarez, R. / Xove, X. (2002): *Gramática galega* (Vigo: Galaxia).
- Beaugrande, R.-A. de / Dressler, W. Ulrich (1997): *Introducción a la lingüística del texto* (Barcelona: Ariel).
- Carballo Calero, R. (1979) [1966]: *Gramática elemental del gallego común* (Vigo: Galaxia).
- Carré Alvarellos, L. (1967): *Gramática gallega* (A Cruña: Editorial Moret).
- Costa Casas, X. X. / González Refoxo, M. A. / Morán Fraga, C. C. / Rábade Castiñeira, X. C. (1988): *Nova gramática para aprendizaxe da lingua* (A Coruña: Vía Láctea).

Freixeiro Mato, X. R. (1998): *Gramática da lingua galega*. I. *Fonética e Fonoloxía*; (1999): II. *Semántica*; (2000): III. *Morfosintaxe*; (2003): Vol. IV. *Gramática do texto* (Vigo: A Nosa Terra).

Freixeiro Mato, X. R. (2001): *Manual de gramática galega* (Vigo: A Nosa Terra).

Lugrís Freire, M. (1931) [1922]: *Gramática do idioma galego* (A Coruña: Moret).

Saco Arce, J. A. (1868): *Gramática Gallega* (Lugo: Imprenta de Soto Freire).

Koch, I. Villaça / Travaglia, L. C. (2002) [1990]: *A Coerência Textual* (São Paulo: Editora Contexto)

Marcos Bagno (1997), *A Língua de Eulália. Novela Sociolingüística*, São Paulo, Editora Contexto.

Marcos Bagno (1999), *Preconceito lingüístico. O que é, como se faz*, São Paulo, Loyola.

As dúas obras do sociolingüista e profesor do Departamento de Lingüística da Universidade de Basíla, Marcos Bagno, que escollemos gardan unha forte relación entre si e unha serie de semellanzas cos preconceptos e tópicos que sobre o galego actúan freando o seu uso, algo que as fai especialmente atractivas para o público do noso país, que pode ver reflectidas en moitas destas consideracións algunhas das dificultades contra as que o noso idioma tivo e ten de combater. Velaí, por exemplo, a necesidade de elevar o grao da propia autoestima lingüística, impóndonos como falantes competentes na nosa lingua ou a posibilidade de todas as variedades lingüísticas serviren de vehículo de expresión para calquera tipo de mensaxe ou discurso se nelas se realizar o mesmo investimento que acompaña a norma padrón das linguas normalizadas consideradas referentes de cultura. Estes dous textos presentan, do noso punto de vista, un indubitábel interese para aquelas persoas preocupadas pola(s) lingua(s) no noso país, que obterán tras a súa lectura, para alén dun coñecemento máis profundo da realidade lingüística brasileira, un suxestivo contraste con moitos dos prexuízos derivados do conflito lingüístico galego e unha aproveitábel reflexión sobre o ensino da lingua propia, en grande medida e por diferentes causas, tamén unha “lingua estranxeira” para moitas alumnas e moitos alumnos.

Permitásenos, antes de afondarmos na descrición destes dous textos, recomendar tamén outras obras máis recentes do autor en que continúa a súa loita particular contra a exclusión lingüística e social das persoas máis desfavorecidas centrándose no ensino como sistema ideolóxico de dominación e de transmisión de valores e crenzas marxinatoras da grande maioría da poboación. Referímonos a *Dramática da língua portuguesa: tradição gramatical, mídia & exclusão social* do ano 2000, *Português ou Brasileiro? (Um convite à pesquisa)* do ano 2001 ou a traballos como a organización do volume *Lingüística da Norma* publicado no 2002. En todos eles, e como unha máxima que percorre toda a súa obra, aparece a idea da necesidade de non cedermos todo o espazo sobre a reflexión lingüística aos especialistas, aos filólogos, unha tese acompañada da pretensión de examinar desde perspectivas diferentes á tradicional a realidade e, sobre todo de a non construír ficticia e verticalmente:

O exame do concepto de norma, mesmo quando seguido do adxectivo lingüística, transpõe os límites dos domínios das ciencias da linguagen, obrigando à sempre inevitável e salutar intersección com outros campos de coñecemento, como a filosofía, a sociología, a antropología, a pedagogía, a historia, para citar apenas os máis evidentes. Esta coletânea, me parece, mostra isso muito bem, sobretudo quando dela participan, além de uma maioria de lingüistas, pesquisadores das ciencias sociais e das ciencias da educación.

Este libro (...) pretende ser uma contribuição para que o debate sobre uma ampla agenda de temas importantes –a variação e mudança lingüísticas, o ensino de língua nas escolas, a formação dos professores de portugueses, o lugar do ensino da gramática, as políticas de letramento e tantos outros– se faça de modo mais claro e com critérios científicos mais coerentes do que os argumentos nitidamente preconceituosos e de perfil marcadamente autoritário que, nos últimos tempos, têm emergido no panorama cultural brasileiro, sobretudo nos meios de comunicação (“Primeiras *palavras*”, *Lingüística da Norma*, p. 10).

Na primeira das obras (*A Língua de Eulália. Novela Sociolingüística*) o sociolingüista brasileiro preséntanos o percurso educativo extraacadémico de tres estudantes moi novas que pasan de criticar a lingua de Eulalia, unha amiga e antiga empregada de Irene, profesora de lingüística na universidade e tía dunha das raparigas, a recibiren leccións sobre o seu falar ampliando o seu coñecemento “escolar” e liberándose dos preconceitos derivados en grande medida dos seus estudos.

A explicación do título da novela deita luz sobre as súas pretensións, pois do que se trata é de describir as características xerais do portugués non padrón falado no Brasil a partir do seu emprego por unha utente dunha das súas variedades, a Eulalia (“a que fala bonito”, “a que fala ben”) procurando eliminar toda a carga prexuízosa e de sinalar a lóxica da súa lingua descrita como máis unha variante sistemática, coherente e funcional. Todo isto é analizado desde a perspectiva de valorar non só o modelo lingüístico das persoas usuarias desta variedade, senón, e fundamentalmente, de aumentar a súa autoestima diminuindo a fronteira abismal entre o coñecemento intelectual e a sabedoría popular, algo que, levado ao mundo do ensino debería polo menos reducir a distancia existente entre o padrón lingüístico e o uso real da lingua por parte dos falantes cultos, abrindo asemade un espazo para outras variedades lingüísticas.

Igual que o título e o nome da principal protagonista (practicamente ausente ou secundarizada no texto, o que resta certa credibilidade ás ousadas propostas democráticas do autor ao non existir coherencia entre praxe e teoría) non son circunstanciais, tampouco parece selo a escolla de personaxes femininas, tanto por representaren a abundancia do sexo feminino na profesión educativa como por exemplificaren o feito de se deixaren levar as mulleres máis facilmente do que os varóns

polos prexuízos, ao seren educadas e asimilaren con maior facilidade a necesidade de se achegaren á norma, a todo aquilo considerado como o correcto e o mellor valorado. Mais existe tamén unha reivindicación implícita ao ceder a palabra para crear o seu propio discurso ás mulleres que, igual que os escravos negros ou os grupos étnicos e estratos sociais desprestixiados, conformaron unha maioría da poboación brasileira silenciada e oprimida fronte a un pequeno colectivo esencialmente masculino, branco e oligárquico. O combate contra a crenza de estarmos perante unha “língua errada” en que non se respectan as normas ditas pola gramática acompáñase, pois, dun continuo acrecentamento de nocións ideolóxicas con que se quere establecer a relación entre a discriminación lingüística e a social e política.

A obra ofrece varios atractivos: en primeiro lugar, e para alén de se dotar dun molde ficcional, preséntase en forma de diálogo, utilizando un modelo pedagóxico moi acaído para persoas non especializadas non tema, pois, como afirma o autor “*todo professor é professor de língua*, já que ele se serve da língua como meio de transmissão dos conteúdos que lhe cabe ensinar” (p. 29); en segundo lugar, dános a posibilidade de nos achegar a través das súas páxinas a unha pequena historia do portugués, ao explicar moitos dos fenómenos existentes nas variedades non padronizadas recorrendo á lingua utilizada noutras etapas e incluíndo en ocasións a comparación cos resultados de certas palabras noutros idiomas románicos; en terceiro lugar, e este é o principal aspecto que queremos salientar, ofrece interesantes exemplos de preconceitos e falsas consideracións sobre a lingua, centrando a existencia de mitos e prexuízos na discriminación de grande parte da poboación, deixando como mensaxe central do seu discurso o necesario recoñecemento de ser a lingua, fundamentalmente, variación:

Na verdade, se quisermos ser exatas e precisas na hora de dar nome a uma língua, teríamos de dizer, por exemplo, falando da Vera: “Esta é a língua portuguesa, falada no Brasil, em 2001, na região Sudeste, no estado e na cidade de São Paulo, por uma mulher branca, de 21 anos, de classe média, professora primária, cursando universidade” etc. Ou seja, teríamos de levar em conta todos os elementos –chamados *variáveis*– que compõem uma *variedade* (p. 20).

O libro divídese en 19 pequenos capítulos que, baixo o fío condutor das clases sobre o portugués non padrón aprendido a través da lingua de Eulália, representan unha serie de temas relacionados coa variación lingüística e a aprendizaxe do portugués no Brasil. Para alén destes apartados, un capítulo introdutor (“A chegada”) e outro final (“A partida”) danlle aparencia de novela a estes apuntamentos dialogados que, xa na derradeira parte da obra e como un apéndice final (“Mais duas palavrinhas e sugestões de leitura”), descubren o seu obxectivo didáctico e o seu destinatario principal nun alumnado cunha certa especialización, mesmo indicando unha serie de recomendacións bibliográficas comentadas.

Preconceito lingüístico. O que é, como se faz combate a confusión entre lingua e gramática normativa debullando nun primeiro apartado titulado “A mitología do preconceito lingüístico” cada un dos grandes mitos en que se alicerza a discriminación e marxinação lingüística, sempre relacionando esta situación coa colonización e exclusión social das persoas que non falan a lingua do poder, ou, dito doutra maneira, fuxindo de abstraccións que fagan da lingua un obxecto de estudo independente do seu contexto social, cultural e político:

Existe um mito ingênuo de que a linguagem humana tem a finalidade de “comunicar”, de “transmitir idéias” –mito que as modernas correntes da lingüística vêm tratando de demolir, provando que a linguagem é muitas vezes um poderoso instrumento de *ocultação da verdade*, de *manipulação do outro*, de *controle*, de *intimidação*, de *opressão*, de *emudecimento*. Ao lado dele, também existe o mito de que a escrita tem o objetivo de “difundir as idéias”. No entanto, uma simples investigação histórica mostra que, em muitos casos, a escrita funcionou, e ainda funciona, com a finalidade oposta: ocultar o saber, reservá-lo a uns poucos para garantir o poder àqueles que a ela têm acesso (p- 133).

A enumeración dos mitos tratados polo autor, con abondosa exemplificación, serve para observar as semellanzas con moitos dos tópicos que xurdiron no proceso de recuperación da nosa lingua como idioma de cultura, cuxa pegada aínda se deixa notar hoxe entre certos sectores da cidadanía galega. Se o primeiro deles (“A lingua portuguesa falada no Brasil apresenta uma unidade surpreendente”) se afasta ou mesmo contradi as ideas sobre o galego, habitualmente descrito como un idioma excesivamente fragmentado pola carencia dunha norma culta de referencia, no resto dos casos, con lixeiras adaptacións contextuais, observamos os mesmos prexuízos existentes nalgún momento da historia máis recente da lingua galega:

Mito nº 2: “Brasileiro não sabe português. Só em Portugal se fala bem português”.

Mito nº 3: “Português é muito difícil”.

Mito nº 4: “As pessoas sem instrução falam tudo errado”.

Mito nº 5: “O lugar onde melhor se fala português no Brasil é o Maranhão”.

Mito nº 6: “O certo é falar assim porque se escreve assim”.

Mito nº 7: “É preciso saber gramática para falar e escrever bem”.

Mito nº 8: “O domínio da norma culta é um instrumento de ascensão social”.

Na segunda parte da obra, “O círculo vicioso do preconceito lingüístico”, introducímonos naqueles elementos que transmitiron e continúan a perpetuar as crenzas a respecto dos mitos anteriormente descritos. Eis agora a “Santíssima Trindade” do preconceito lingüístico: a gramática tradicional, os métodos tradicionais de ensino e os libros didácticos.

Após unha exemplificación dos erros e falsos valores ensinados nas escolas e difundidos nos medios de comunicación a través do que o autor denomina como “comandos paragramaticais”, ou instrumentos de reiteración e perpetuación dos falsos mitos sobre a lingua, o autor centra xa a terceira parte da obra na súa propia proposta de combate contra os preconceitos. Coa “deconstrución do preconceito lingüístico”, e tras un primeiro exercicio de recoñecemento do problema (velai outra idea chave: a necesidade de facermos visíbel e sermos conscientes da doenza para a podermos cortar) Bagno insiste na mudanza de actitudes como paso previo a calquera solución. A partir de aquí todo resulta máis doado: o profesorado debe (in)formarse, facer unha crítica activa da súa posta en práctica na aula e asumir que a ciencia da linguaxe evolúe, para, nunha especie de decálogo final, recoller as DEZ-CISÕES (p. 142-145) para un ensino da lingua non (ou menos) preconceituoso. Seguindo a súa simboloxía relixiosa, estas dez cuestións poderían resumirse en dúas: recoñecer as persoas usuarias dunha lingua como competentes nesa lingua sen confundir erro lingüístico con erro ortográfico e observar a lingua como un organismo en constante evolución.

O último capítulo do libro ten un carácter máis circunstancial ao se ligar coas polémicas aparecidas na prensa entre partidarios e combatentes do ensino da Gramática tradicional. A defensa das posturas do autor pretende evidenciar que as xustificacións dalgunhas autoridades ao recriminaren estes posicionamentos modernos está baseada na perda duns valores que sustentaban, en parte, os seus privilexios.

A compatibilidade do presuposto da variación como fundamental para a descrición lingüística coa afirmación de termos galegos e galegas a sorte de contar cunha importante proxección e comunicación co mundo lusófono, negada por Dasilva (2003: 124), aparece en certas indicacións que poderían considerarse marxinais, pois, evidentemente, non son o centro da obra, cuxo eixe descansa sobre a necesidade de mudar o sistema de ensino da lingua no Brasil, adecuándose ao contexto en que este se insire. Xustamente estamos a falar da adecuación e a aceptabilidade que no apartado 5 do capítulo 3º (“Então vale tudo?”) se recolle de maneira maxistral:

Uma das principais tarefas do professor de língua é conscientizar seu aluno de que a língua é como um grande guarda-roupa, onde é possível encontrar todo tipo de vestimenta. Ninguém vai só de maiô fazer compras num shopping-center, nem vai entrar na praia, num dia de sol quente, usando terno de lã, chapéu de feltro e luvas... (p. 130).

Mais tamén na novela de aprendizaxe *A língua de Eulália*, onde as tres transformadas estudantes voltan ás súas casas completamente diferentes de como chegaron, o autor enfatiza a idea da uniformidade lingüística entre o portugués padrón e o portugués non padrón, insistindo en que as semellanzas entre eles son moito

maiores que as diferenzas, feito que explica a intercomunicación entre un habitante analfabeto das palafitas do Amazonas e un falante escolarizado do Río Grande do Sul. Non existe, pois, contradición entre o recoñecemento da diversidade e da variación lingüística (“Como sempre, tudo vai depender de *quem diz o quê, a quen, como, quando, onde, por quê e visando que efeito...*”, *Preconceito lingüístico*, p. 131) coa aposta por realizar unha política lingüística e cultural de aproximación aos países da lusofonía onde achamos, para alén de resolucións prácticas de grande interese no campo da neoloxía e da terminoloxía, lecturas que poden suplir as nosas carencias e axudarnos a equilibrar a tradicional dependencia da realidade cultural española. Dito con outras palabras, a comunicación coa lusofonía e un maior coñecemento da súa realidade lingüística e cultural son, adecuadamente aproveitadas, instrumentos de grande utilidade para frearmos ou reducirmos os efectos do acelerado proceso de substitución lingüística que está a padecer Galiza.

Goretti Sanmartín Rei
goretti@udc.es

Referencias bibliográficas

- Bagno, M. (2000): *Dramática da língua portuguesa: tradição gramatical, mídia & exclusão social* (São Paulo: Loyola).
- Bagno, M. (2001): *Português ou Brasileiro? (Um convite à pesquisa)* (São Paulo: Parábola).
- Bagno, M. (org.) (2002): *Lingüística da norma* (São Paulo: Loyola).
- Dasilva, X. M. (2003): “O efecto Marcos Bagno na lingua do Brasil”, *Grial*, 160: 123-124.