

## LIMIAR

Recentemente o Instituto da Lingua Galega quixo facer unha instántanea do estado actual da lexicografía galega co fin de, a partir dela, propiciar unha reflexión e planificar o traballo da súa sección de lexicografía para os próximos anos. Con tal motivo convidou a participar a distintos especialistas, da casa e de fóra, e encargoulles a realización de traballos orixinais nos que desen conta dalgún aspecto da súa propia investigación que fose relevante neste eido e, deste xeito, contribuísen a fixar as áreas temáticas preferentes e delimitar mellor, dentro delas, as vías polas que haxa que encamiñar no futuro os esforzos lexicográficos.

Neste volume reúnen 19 contribucións, mais poderían figurar algunhas máis que, estando previstas no chamamento inicial, non chegaron a termo. A uns e a outros contribuíntes exprésalles o Instituto da Lingua Galega o seu agradecemento pola colaboración prestada.

\*\*\*

A idea inicial foi abordar de forma proporcional tres grandes perspectivas: (a) a historiografía e a metodoloxía lexicográficas, (b) os recursos lexicográficos no presente e as súas perspectivas de futuro e (c) a lexicografía académica. Ao cabo as seccións non resultaron exactamente proporcionais mais nas tres comparecen traballos representativos e de innegable valor.

Abre a sección de metodoloxía o traballo de Manuel Alvar Ezquerro (Universidad Complutense de Madrid). Nel dá noticia pormenorizada dos aspectos máis cotiáns e íntimos do proceso de construción do monumental –11 volumes e 10.522 páxinas– e recentemente aparecido (2007) *Nuevo tesoro lexicográfico del español (s. XIV-1726) (NTLE)*, feito en colaboración con Lidio Nieto Jiménez. Este capítulo é un bo exemplo de “caderno de

bitácora” da fabricación dun dicionario de dicionarios. Nel non faltan as alusións aos problemas do lexicógrafo cunha informática balbuciente (finais dos 80 e comezos dos 90), ao compromiso de seleccionar os materiais nunha longa lista de fontes, ao intenso esforzo de lematización de textos dunha elevadísima variación gráfica e, tamén, á continxencia de ter que enfrontar unha reforma ortográfica na metade do camiño (RAE 2000). No segundo capítulo, Dieter Messner (Paris-Lodron-Universität Salzburg) explora no seu tamén monumental e aínda en curso de publicación *Dicionário dos dicionários portugueses* (1994- ) (20 volumes e 12.000 páxinas) o aproveitamento, non sempre –ou case nunca– recoñecido, da información lexicográfica dos dicionarios na produción dicionarística portuguesa ata 1858. Nela sobresaen como fontes lexicográficas, en distintas épocas, os propios dicionarios portugueses (nun lugar destacado o Bluteau (1712-1728), os españois (moi especialmente o *Tesoro de la lengua castellana o española* de Covarrubias (1611) e o *Diccionario de Autoridades* da RAE (1726-1739) e, en menor medida, os franceses (a *Encyclopédie* (1752). D. Messner é especialmente crítico co abuso deste procedemento no *Diccionario da Lingoa Portuguesa publicado pela Academia Real das Sciencias de Lisboa* (1793), que considera un exemplo reprobable dos “dicionários que copiam por inteiro o que está em outros dicionários” (neste caso o *Diccionario de Autoridades*). Xosé-Álvaro Porto Dapena (Universidade da Coruña) ofrece no terceiro capítulo o seu punto de vista sobre o papel e o lugar da información gramatical nos dicionarios xerais e comúns. Desde a súa longa experiencia como lexicógrafo e gramático e moi especialmente desde a praxe actual da elaboración, como director, do *Diccionario “Coruña” de la lengua española actual (DCLEA)*, de futura aparición, revisa os aspectos teóricos da distinción entre Léxico e Gramática e o espiñento cometido de establecer na información gramatical os límites entre o que é obxecto da gramática e o que é obxecto do dicionario. Pasa revista a practicamente todos os valores gramaticais susceptibles de figuraren nun dicionario, tanto na macroestrutura como na microestrutura. Por último, Telmo Verdelho (Universidade de Aveiro) fai no seu capítulo unha historia crítica da lexicografía portuguesa ata 1900, desprovista cronicamente do amparo das institucións políticas nacionais. Entre as súas características menciona unha aparición tardía –comparativamente con outras: italiana, española, francesa– e a orixinalidade de iniciarse, en dicionarios bilingües, tomando como punto de partida o vernáculo e non o latín. Destaca nela, como fíos condutores sucesivos, a liña de obras nadas como resposta ao encontro interlingüístico con Asia e América (o Brasil) –na que foi verdadeira pioneira e na que tiveron un papel destacadísimo os Xesuítas (ata

1750)–, e como resposta ao confronto moderno do portugués coas linguas de Europa. A nivel de autores salienta a figura do Rafael Bluteau e o seu monumental *Vocabulario Portuguez...* (1712-1728).

O segundo bloque de contribucións reúne traballos dedicados á presentación de proxectos e recursos de aproveitamento lexicográfico desenvolvidos no ámbito do galego, co engadido dalgún do portugués. Un deles, o capítulo 5, de María Xesús Bugarín (ILG), ten carácter individual e preséntase como investigación concluída. Nel explica a autora a elaboración da súa obra *O galego fundamental. Diccionario de frecuencias* (2007), referida ao galego moderno. Detense na indicación das utilidades varias dunha obra deste tipo –é, neste dominio, a primeira e única publicación sobre o galego– e, sobre todo, relata con abundancia de exemplos ilustrativos o proceso de elaboración (construción do corpus, tratamento dos materiais e obtención de resultados).

Nos outros capítulos deste segundo bloque dáse noticia de proxectos, recursos e investigacións lexicográficas en curso, realizados colectivamente por equipos de investigación de carácter universitario ou institucional. Coinciden todos eles en presentaren experiencias, propostas e logros na lingüística de corpus ou na fabricación de bases de datos de orientación lexicográfica. No capítulo 6 describe Eva Domínguez o *Corpus de Referencia do Galego Actual (CORGA)* do Centro “Ramón Piñeiro” para a Investigación en Humanidades. Centra a atención, sobre todo, no etiquetador e lematizador do corpus, do que forman parte o sistema de etiquetas ou *tagset*, o lexicón, o preprocesador (sen resolución de ambigüidades) e o etiquetador (resolución de ambigüidades). No capítulo 7 expoñen Francisco Fernández Rei e Carme Hermida Gulías (ILG) a estrutura e composición do *Arquivo do Galego Oral (AGO)* do Instituto da Lingua Galega (seis subarquivos, máis de 2.500 horas de gravacións, voces de 7.000 informantes e 315 concellos de Galicia + 38 da Galicia exterior + mostras dos concellos cacereños do Val de Xálima) e ilustran cun caso práctico o aproveitamento léxico dos seus materiais. No capítulo 8 Xesús Ferro Ruibal, director do proxecto *Fraseoloxía Galega* do Centro “Ramón Piñeiro”, rescata das compilacións paremiolóxicas de Vicente Llopiz (1954a e 1954b) e de Francisco Vázquez Saco (1962) unha mancha de valores semánticos, variantes gramaticais infrecuentes, voces xa pouco frecuentes ou hibernadas e mesmo voces íntegras que non aparecen rexistrados nas obras lexicográficas compiladas no *Diccionario de diccionarios* de Antón Santamarina (2003b), o que demostra que “a fraseoloxía [...] aínda é unha despensa léxica” e que hai que tela máis presente. No capítulo 9 Ana Belén Crespo Bastos e Xosé María Gómez Clemente (Universidade de Vigo) presentan os novos criterios de identificación de

neoloxismos empregados nos traballos do Observatorio de Neoloxía da Universidade de Vigo. Inciden, unha vez máis, na necesidade de revisión do concepto de “neoloxismo”, na importancia de establecer con coherencia as marcas dos dicionarios de neoloxismos (marcas temporais, marcas de frecuencia absoluta e dispersións temporal e marcas lingüísticas) e na importancia das tecnoloxías informáticas como ferramenta auxiliar (extractores semiautomáticos de neoloxismos...). Xavier Gómez Guinovart, director do Seminario de Lingüística Informática (SLI) da Universidade de Vigo, ocúpase no capítulo 10 dos logros lexicográficos e terminolóxicos obtidos a partir dos corpóra da Universidade de Vigo creados polo SLI e o OdN. Explica as características dos corpóra *Corpus Lingüístico da Universidade de Vigo (CLUVI)* e *Corpus Técnico do Galego (CTG)*, e, como produtos propiamente lexicográficos, presenta o *Dicionario CLUVI Inglés-Galego* e o *Banco de Datos Terminolóxico da Universidade de Vigo*. Detense nos proxectos que o grupo ten en perspectiva e entre eles destaca a colaboración co Instituto da Lingua Galega en distintas empresas científicas. No seguinte capítulo, o 11, Paulo Martínez Lema (ILG) amosa a nova versión do *Tesouro Medieval Informatizado da Lingua Galega (2007)*, o corpus lingüístico medieval do Instituto da Lingua Galega (USC). Explica a súa estrutura e características (tipoloxía textual, cronoloxía e glotoloxía, microestrutura do corpus) e indica as expectativas que ten abertas como banco de datos doutros proxectos lexicográficos (*Dicionario Medieval Galego-Portugués*, *Dicionario Histórico Galego e Inventario Toponímico da Galicia Medieval*). No capítulo 13 Xusto Rodríguez Río (USC) describe o *bUSCatermos* e mais o *Banco de Datos Terminolóxico (BDT)*, dúas instancias dun mesmo recurso terminolóxico do Servizo de Normalización Lingüística da Universidade de Santiago de Compostela. Explica moi sucintamente a composición das fontes empregadas (propias e alleas) e o proceso de reelaboración do material recollido (aplicación do enfoque onomasiolóxico, arrequentamento da información das fontes e análise valorativa das propostas alleas). No capítulo 13, Maria Francisca Xavier (Centro de Linguística da Universidade Nova de Lisboa) ofrece, como directora, unha panorámica de dous proxectos de investigación lexicográfica do portugués medieval: o *Dicionário do Português Medieval (DPM)* e o *Corpus Informatizado do Português Medieval (CIPM)*. En relación co primeiro, presenta as publicacións derivadas, explica o sistema de traballo (preparación dos datos, identificación das unidades léxicas e análise delas) e explicita a estrutura da ficha lexicográfica (*vedetta*, clasificación das palabras, etimoloxía, acepcións, aboamentos e remisións). En relación co corpus, ofrece a listaxe de obras procesadas.

O terceiro bloque acolle os seis traballos recibidos de representantes autorizados de Academias e institucións investidas de potestade na fixación do estándar lingüístico das súas respectivas comunidades ou altamente representativas delas. Algúns, seguindo as indicacións dos editores deste volume, preferiron ofrecer unha visión histórica da actividade lexicográfica da institución representada, outros optaron por ocuparse dun aspecto ou obra concreta e outros decidiron conxugar as dúas solucións anteriores. No nome da Real Academia Galega e como Secretario dela, Manuel González presenta no capítulo 14 a actividade dicionarística do Seminario de Lexicografía desta institución –do que é director– nos últimos vinte anos. Despois de pasar unha rápida revista ao *Diccionario da lingua galega* (ILG/RAG 1990) e ao *Diccionario da Real Academia Galega* (RAG 1997), expón as características da futura versión da segunda das obras mencionadas, que conterà 50.000 entradas. Detense nas fontes empregadas (obras lexicográficas anteriores, corpus representativo da lingua escrita do galego moderno e corpus representativo da lingua oral actual), na macroestrutura da obra e na súa microestrutura. Joaquim Rafel i Fontanals comparece, no capítulo 15, na representación do Institut d’ Estudis Catalans. No plano xeral, ofrece unha visión da historia, das funcións e da estruturación en seccións do Institut. En referencia aos traballos lexicográficos da Sección Filolóxica, dá noticia dos dicionarios de 1995 e 2001 e detense na elaboración do *Diccionari del català contemporani* (DCC). Presentao como produto *in fieri* da realización, en dúas fases, de proxectos previos: o *Corpus Textual Informatizat de la Llengua Catalana* (CTILC) e a *Base de Dades Lexicogràfica* (BDLex). Partindo da información contida nestes dous recursos e contando cunha excelente “estación de traballo lexicográfica”, estase a realizar na actualidade este dicionario descritivo, “que tiene por objeto la definición de las unidades léxicas de la lengua desde el punto de vista de su contenido y de su utilización real, sin restricciones basadas en criterios prescriptivos”. No capítulo 16, João Malaca Casteleiro (Academia das Ciéncias de Lisboa) fai unha síntese da historia dos dicionarios da Academia desde o ano 1793, no que aparece o *Diccionario da lingua portuguesa*, ata a actualidade, cando xa está concluído, aínda que inédito –desde o 2005–, o *Novo Dicionário da Língua Portuguesa*. Ao seu antecedente, o *Dicionário da Língua Portuguesa Contemporânea* (2001), dedícalle boa parte do capítulo. Explica o ámbito cronolóxico (séculos XIX e XX), o volume da súa información (70.000 entradas e 170.000 significados e acepcións) e o seu carácter panluso (presenza notable de brasileirismos, africanismos e asiaticismos). Por parte da Real Academia Española e como Secretario dela, Guillermo Rojo (USC) fai no capítulo 17

un balance histórico da actividade lexicográfica da RAE. Correspondendo co pasado, explica as características básicas das sucesivas edicións do *DRAE* anteriores á de 2001. Como actividades presentes da Academia destaca a realización dos córpora moderno (*CREA*) e diacrónico (*CORDE*), culminada coa publicación, en 2001, do *Diccionario de la lengua española* (edición en CD-ROM en 2003). De cara ao futuro sinala, como filosofía de traballo, a importancia da colaboración entre todas as Academias e, no terreo das obras lexicográficas, a posta en marcha do *Nuevo diccionario histórico de la lengua española*, do *Diccionario académico de americanismos* e a próxima publicación dun novo *DRAE* (2013?) no que xa non se inclúa a información diacrónica e diatópica correspondente aos dous dicionarios mencionados. No capítulo 18 e na representación da Opera del Vocabolario Italiano (OVI), Pär Larson describe o marco institucional e filolóxico de OVI (representa a continuación da Accademia della Crusca) e céntrase no *Tesoro della Lingua Italiana delle Origini* (*TLIO*). Realiza unha “visita guiada” a este dicionario do italiano antigo, incluíndo referencias mínimas ao corpus no que se fundamenta. Pasa revista aos oito –en realidade nove– apartados que compoñen a estrutura das entradas (lema –o moderno–, variantes gráficas, etimoloxía, atestacións, observacións lingüísticas, anotacións, valores semánticos estruturados e información sobre os redactores) e destaca tres opcións deste recurso *on-line*: a posibilidade de recuperación de explicacións semánticas nos propios textos (opción *Gloss.*), a indicación do grao de fiabilidade e representatividade lingüística das fontes textuais e a posibilidade de recuperación conxunta das unidades polirremáticas.

No capítulo que clausura o volume, Antón Santamarina (director da Sección de Lexicografía do Instituto da Lingua Galega-USC e membro numerario da Real Academia Galega) contribúe cun extenso texto preparatorio do *Diccionario Histórico da Lingua Galega*. Nas súas páxinas condensa a experiencia e o esforzo de varias décadas de traballo lexicográfico orientado á realización desta vasta obra que situará o galego no restrito grupo de linguas románicas que contén un dicionario histórico. Por varias veces refírese a estes dicionarios como monumentos imprescindibles das culturas nacionais, pero ao que con máis énfase lle atribúe esta cualidade é ao *Oxford English Dictionary* (*OED*), que chega a cualificar como “un dos principais monumentos da cultura mundial do século XX”. Nunha primeira parte de carácter historiográfico bota unha ollada aos máis importantes dicionarios históricos. Cita algúns antecedentes antigos e detense máis demoradamente nos dicionarios de linguas da Europa occidental, moi maioritariamente de linguas xermánicas (alemán,



inglés, holandés, afrikaans, sueco, danés, noruegués), aínda que tamén dalgunha lingua románica (francés e, moi sinaladamente, catalán). Na segunda parte vai debullando os trazos constitutivos do futuro *DHG*: o alcance cronolóxico ou profundidade histórica, as fontes textuais (córpora e dicionarios), as características da macroestrutura e a organización da microestrutura (información fonética, variantes gráficas, etimoloxía, información léxico-semántica). Por último valora a viabilidade da empresa atendendo a cinco parámetros: (a) quen debe levala a cabo, (b) en que institución, (c) canto tempo vaia levar, (d) custo das actividades e (e) como se vaia facer (tecnoloxías informáticas).

\*\*\*

No nome do Instituto da Lingua Galega queremos deixar constancia do agradecemento a distintas institucións e organismos polas axudas recibidas. A Real Academia Galega patrocinou, ofreceu axuda económica e deu acollida na súa sede a algúns actos. O Ministerio de Educación e Ciencia contribuíu con axuda económica. A Facultade de Filoloxía (USC) deu, como sempre, acollida á maior parte dos actos e eventos a que deu lugar esta iniciativa (realización dun simposio en novembro de 2006). Mención especial merece o Consello da Cultura Galega, que estende o seu patrocinio ao ILG desde o ano 2001, no que inaugurou esta prestixiosa Serie Lingüística, dentro da Colección Base, editada en formato papel e electrónico (<http://www.consellodacultura.org>), este último cunha difusión extraordinaria. Quedamos en débeda tamén cos compañeiros do Instituto da Lingua Galega polo apoio constante ofrecido aos editores durante este tempo, moi especialmente a Paulo Martínez Lema, que, ademais de participar como contribuínte, colaborou eficazmente nas actividades organizativas. Finalmente reiteramos o noso recoñecemento aos autores dos traballos incluídos neste libro, na seguranza de que a súa difusión e lectura fomentará a perseverancia dos lexicógrafos actuais e a incorporación de novos investigadores que aseguren a continuidade no cultivo da lexicografía.

*Os editores*