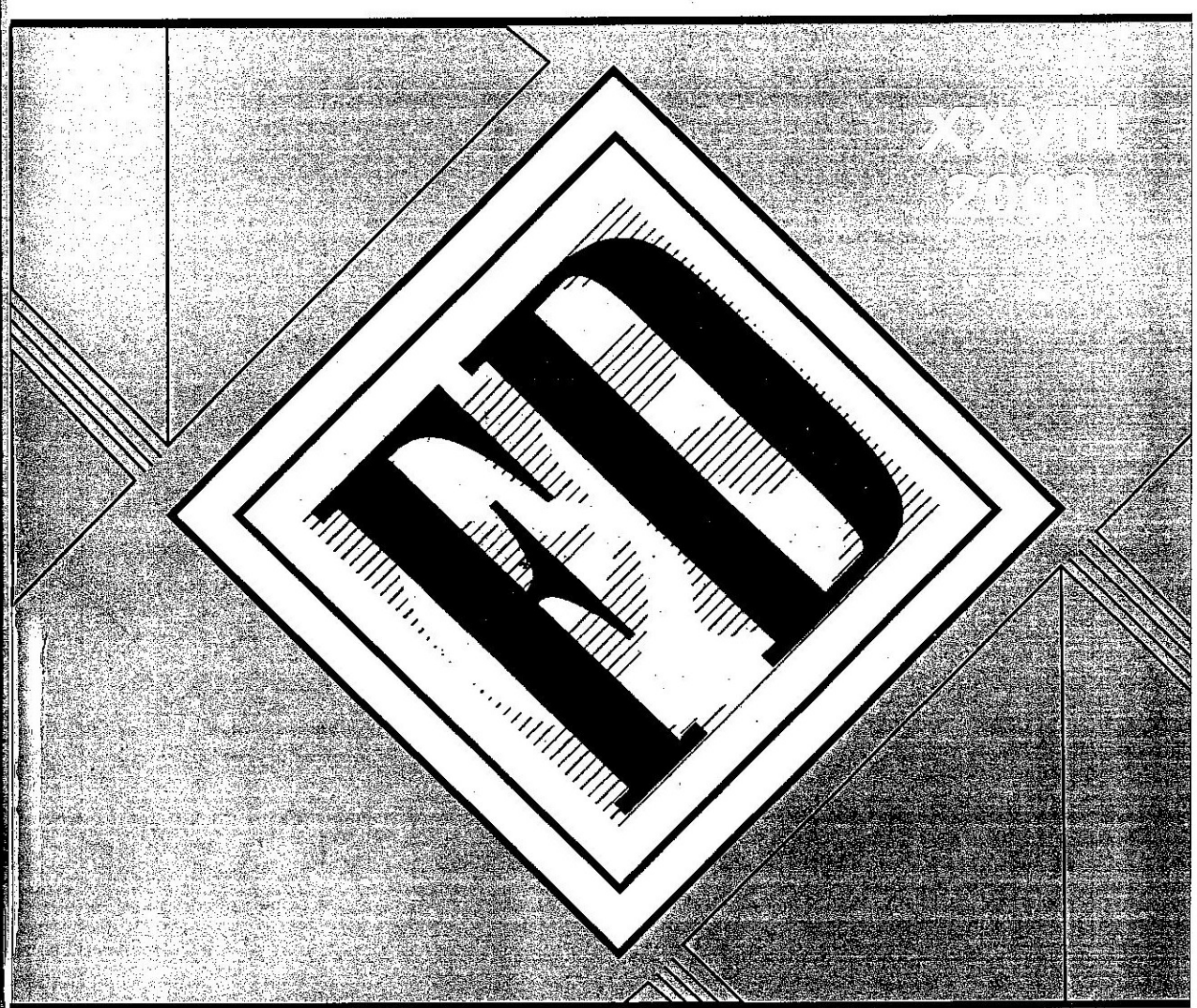
P560

ACADEMIA ROMÂNĂ
INSTITUTUL DE LINGVISTICĂ
„IORGU IORDAN – AL. ROSETTI"

D16171

FONETICĂ ȘI DIALECTOLOGIE



EDITURA ACADEMIEI ROMÂNE

COMITETUL DE REDACȚIE

NICOLAE SARAMANDU – redactor-şef
MARIA MARIN – redactor-şef adjunct
MARIA MĂRDĂRESCU-TEODORESCU – secretar de redacție
MANUELA NEVACI – secretar adjunct de redacție
ANDREI AVRAM
GRIGORE BÂRCUŞ, membru corespondent al Academiei Române
Acad. MATILDA CARAGIU MARIOȚEANU
Acad. MARIUS SALA
TEOFIL TEAHA
DANA-MIHAELA ZAMFIR

Membri din străinătate:
GHEORGHE CARAGEANI (Italia), MICHEL CONTINI (Franța), LARS-ERIK EDLUND (Suedia), MANUEL GONZÁLEZ GONZÁLEZ (Spania), LORENZO MASSOBRIO (Italia), WOLFGANG VIERECK (Germania).

Redactor: ADRIANA GRECU
Tehnoredactor: MARIANA IONICĂ
Coperta: MARIANA ŞERBĂNESCU

Apare o dată pe an.

Toate comenzile pentru lucrările apărute la Editura Academiei Române se vor adresa la:

EDITURA ACADEMIEI ROMÂNE, Calea 13 Septembrie nr. 13, sect. 5, 050711, Bucureşti, România; Tel. 4021–318 81 06/2116, 4021–318 81 46; Fax 4021–318 24 44; E-mail: edacad@ear.ro; Adresa web: www.ear.ro
RODIPET S.A., Piața Presei Libere nr. 1, sect. 1, P.O. Box. 33–57, Bucureşti, România; Tel. 4021-618 51 03; 4021–222 41 26; Fax 4021–222 64 07; E-mail: rodipet@rodipet.ro
ORION PRESS IMPEX 2000 S.R.L., P.O. Box 77–19, sector 3, Bucureşti, România; Tel./Fax: 4021-610 67 65; 4021-210 67 87, Tel.: 0311 044 668; E-mail: office@orionpress.ro

Adresa redacției:

INSTITUTUL DE LINGVISTICĂ
„IORGU IORDAN – AL. ROSETTI"
Calea 13 Septembrie nr. 13, sector 5
050711 Bucureşti, România
Tel. 4021-318 81 06/2207
Tel./fax 4021-318 24 43

Adresa editurii:

EDITURA ACADEMIEI ROMÂNE
Calea 13 Septembrie nr. 13, sector 5
050711, Bucureşti, România
Tel. 4021-318 81 06/2116;
4021-318 81 46; Fax: 4021-318 24 44
E-mail: edacad@ear.ro; Adresa web: www.ear.ro

© 2010, EDITURA ACADEMIEI ROMÂNE

2560

FONETICĂ ŞI DIALECTOLOGI

XXVIII. 2009 Apare o dată

SUMAR

ROSARIO ÁLVAREZ, XOSÉ AFONSO ÁLVAREZ PÉREZ, JOÃO SARAMAGO, XULIO SOUSA FERNÁNDEZ, Presentación de un corpus digital de léxico tradicional: el *Tesouro do léxico patrimonial galego e português* .....
ANDREI AVRAM, Fonetica graiului țigănesc din Ştefăneşti (jud. Ilfov) .....
PARASCHIVA BOBOC, Observații asupra unor sensuri ale cuvântului *haiduc* .....
FEDERICA CUGNO, Novità tecniche e scientifiche, stato dei lavori, risultati e prospettive di alcune imprese geolinguistiche italiane .....
MARIA MARIN, Prepozițiile *dip(t)* şi *dâr(t)* în graiurile dacoromâne .....
IULIA MĂRGĂRIT, Particularitățile unui formant la nivelul graiurilor muntenești: *-eală* .....
DORU MIHĂESCU, Termeni din vocabularul culturii populare a catolicilor din Moldova .....
ECATERINA MIHĂILĂ, Nume de locuri provenite din entopice de origine latină .....
MIHAELA-MARIANA MORCOV, Cuvântul latinesc *bestia* şi descendenții săi romanici .....
LURDES DE CASTRO MOUTINHO, ROSA LÍDIA COIMBRA, ANA MARGARIDA VAZ, Variantes prosódicas do português europeu: o barlavento e o sotavento algarvio
VIRGIL NESTORESCU, Un manual şcolar cu particularități dialectale .....
MANUELA NEVACI, *Cumătră* – pe baza ALRR. *Sinteză* .....
TEOFIL TEAHA, Din lexicul latin moştenit în graiurile românești actuale (V) .....
OHARA DONOVETSKY, Forme şi valori ale verbului în graiurile muntenești (II) .....

NOTE ETIMOLOGICE LEXICALE

IULIA MĂRGĂRIT, Note lexicale şi semantice .....

RECENZII ŞI NOTE BIBLIOGRAFICE

Людвиг Селимски, *Фамилни имена от Северозападна България. Влашки елемент*, Katowice, 2006, 183 p. (*Iulia Mărgărit*) .....
Iulia Mărgărit, *Comentarii etimologice şi semantice. Note şi articole*, Bucureşti, Editura Academiei Române, 2007, 400 p. („Etymologica" 27) (*Daniela Răuțu*) .....

CRONICĂ

The 42nd ALE Editorial Board Meeting, Glasgow (June 25th – 27th, 2009) (*Carmen-Ioana Radu, Manuela Nevaci*) .....
Cercetări dialectale la românii din Voivodina (Serbia) (*Maria Marin, Daniela Răuțu*) .....

INDICE .....

21641

FD, XXVIII, Bucureşti, 2009, p. 1–258

FONETICĂ ŞI DIALECTOLOGIE

XXVIII, 2009 Paraît une fois par an

SOMMAIRE / SUMMARY

ROSARIO ÁLVAREZ, XOSÉ AFONSO ÁLVAREZ PÉREZ, JOÃO SARAMAGO, XULIO SOUSA FERNÁNDEZ, Presentación de un corpus digital de léxico tradicional: el *Tesouro do léxico patrimonial galego e portugués* 5
ANDREI AVRAM, Phonétique du parler tsigane de Ştefăneşti (dép. d'Ilfov) 20
PARASCHIVA BOBOC, Observations on some meanings of the word *haiduc* 40
FEDERICA CUGNO, New technical and scientifical progress. Results and prospects of some geolinguistiques Italian researches 49
MARIA MARIN, The prepositions *dip(t)* and *dâr(t)* in the Dacoromanian idioms 64
IULIA MĂRGĂRIT, Les particularités d'un formant dans les parlers de Munténie: *-ealã* 72
DORU MIHĂESCU, Mots appartenant au vocabulaire de la culture populaire des catholiques de Moldavie 86
ECATERINA MIHĂILĂ, Places names resulted from entopics of Latin origin 111
MIHAELA-MARIANA MORCOV, Lat. *bestia* dans les langues romanes 129
LURDES DE CASTRO MOUTINHO, ROSA LÍDIA COIMBRA, ANA MARGARIDA VAZ, Variantes prosodiques du portugais européen: le barlavento et le sotavento de l'Algarve 142
VIRGIL NESTORESCU, Un manuel scolaire avec des particularités dialectales 152
MANUELA NEVACI, *Cumătră* 'God –sister' – as shown in ALRR. *Synthesis* 159
TEOFIL TEAHA, Mots hérités du latín dans les parlers roumains actuels (V) 164
OHARA DONOVETSKY, Forms and values of the verbs in the Muntenia dialect (II) 183

NOTES ETIMOLOGIQUES ET NOTES LEXICALES

IULIA MĂRGĂRIT, Notes lexicales et notes sémantiques 207

COMPTES RENDUS ET NOTES BIBLIOGRAPHIQUES

Людвиг Селимски, *Фамилни имена от Северозападна България. Влашки елемент*, Katowice, 2006, 183 p. (*Iulia Mărgărit*) 222
Iulia Mărgărit, *Comentarii etimologice şi semantice. Note şi articole*, Bucureşti, Editura Academiei Române, 2007, 400 p. („Etymologica" 27) (*Daniela Răuţu*) 225

CHRONIQUE

The 42nd ALE Editorial Board Meeting, Glasgow (June 25th – 27th, 2009) (*Carmen-Ioana Radu, Manuela Nevaci*) 228
Dialectal researches in the Romanian villages from Vojvodina (Serbia) (*Maria Marin, Daniela* Răuţu) 230

INDEX 232

PRESENTACIÓN DE UN CORPUS DIGITAL DE LÉXICO TRADICIONAL: EL *TESOURO DO LÉXICO PATRIMONIAL GALEGO E PORTUGUÉS*

ROSARIO ÁLVAREZ*, XOSÉ AFONSO ÁLVAREZ PÉREZ**,
JOÃO SARAMAGO**, XULIO SOUSA FERNÁNDEZ*

El proyecto *Tesouro do léxico patrimonial galego e portugués* tiene c objetivo reunir y poner a disposición de la comunidad científica en una ú plataforma digital diferentes materiales lexicográficos con referencia geográ procedentes de fuentes diversas, tanto publicadas como inéditas, y en heteroge forma de conservación. Esta iniciativa pretende contribuir a conservar y pon disposición de la comunidad científica esta importante parte del patrim inmaterial de nuestros países, que en muchos casos está amenazada po abandono de las formas de vida tradicionales. El *Tesouro* es un proy internacional coordinado por el Instituto da Lingua Galega (Universidade Santiago de Compostela) y en él participan investigadores gallegos, portugues brasileños, responsables del material incorporado desde sus áreas respectivas *Tesouro* tendrá acceso universal y gratuito a través de Internet en una base de d que permitirá consultas por lema o variante y que incorporará una aplicación la cartografía automática de los resultados seleccionados.

1. OBJETIVOS

Las lenguas son un patrimonio inmaterial que hay que preservar. cambios sociales y tecnológicos producidos en las últimas décadas, con la pér del mundo rural tradicional y sus modos de vida, traen consigo el desuso de mu conocimiento léxico y, en consecuencia, el abandono y la extinción de ciento palabras patrimoniales. A estas pérdidas no escapan lenguas asentadas y reconoc en sus sociedades, como el castellano o el portugués, y aun menos aquellas como el gallego, tienen que vivir en situaciones de extrema dificultad. El *Tesc*

* Universidade de Santiago de Compostela.
** Centro de Linguistica de Universidade de Lisboa.

FD, XXVIII, Bucureşti, 2009, p. 5–19

tiene como fin declarado documentar la forma, el significado y la localización geográfica de esas voces que pertenecen a un modo de vida ya desaparecido o en trance de extinguirse.

El objetivo básico de este proyecto es, como ya se ha señalado, integrar en un corpus único el abundante material léxico dialectal disperso en obras y soportes muy heterogéneos (monografías dialectales, atlas, cuadernos de campo, bases de datos, trabajos onomasiológicos, etc.). Además de facilitar la consulta de estos materiales, esta herramienta tiene el objetivo de hacer visible y preservar la información contenida en una gran cantidad de trabajos académicos que en su mayor parte permanecen inéditos y resultan de muy difícil consulta; con el *Tesouro* se facilita el acceso universal a una información muy rica y valiosa, hasta ahora oculta para la mayoría de los investigadores.

En el *Tesouro* se archivan y procesan todas las voces registradas, junto con la información lexicográfica, etnográfica y bibliográfica correspondiente, además de la localización geográfica de cada forma. Para ello, se marca como un objetivo principal la creación de una base de datos dialectal que permita almacenar y consultar información léxica de las variedades dialectales habladas en Galicia, Portugal y Brasil. Esta base está lematizada, como se explicará más adelante, por lo que el usuario puede obtener toda la información lexicográfica haciendo consultas por un lema que conoce (p. ej., port. *igreja* ou gal. *igrexa*) o por una variante (p. ej., gal. *eirexa*). El otro objetivo principal del *Tesouro* es ofrecer la cartografía con la distribución territorial de los lemas o sus variantes, dentro del dominio gallego y del dominio portugués, en Portugal y Brasil.

El cumplimiento de estos objetivos conlleva la redacción de un protocolo metodológico de actuación lingüística (lemas, variantes, clasificación semántica, etc.) y el desarrollo de las herramientas informáticas precisas para la colección, archivo y tratamiento de toda la heterogeneidad de los datos, y para el diseño de un sistema de búsquedas eficiente. Desde el punto de vista informático, el objetivo es conseguir una herramienta con una *interface* de fácil manejo y accesible para el usuario. El diseño deberá superar diferentes desafíos, especialmente el almacenamiento, integración y recuperación de materiales con formato y origen muy diferentes y pertenecientes a tres subsistemas lingüísticos distintos. Otra característica que requerirá especial atención será la cartografía lingüística automatizada, según las necesidades del usuario y a partir de las búsquedas por él realizadas. Tanto la metodología lingüística como el diseño informático serán ensayados sobre nuestro corpus y aplicados a él, pero deben ser aprovechables en otros dominios lingüísticos.

El conjunto de las fuentes está constituido sobre todo por formas léxicas con sus significados, pero en general no faltan muestras de lengua viva que las contienen, esto es, ejemplos, refranes, canciones, etc. Es también frecuente que estos materiales lexicográficos se acompañen de fotografías o dibujos, que reproducen y preservan para la posteridad lugares, ritos, costumbres, construcciones,

instrumentos y todo tipo de objetos. El *Tesouro* acoge y trata adecuadamente e datos y materiales complementarios, para que siempre estén a disposición del usuar

En fin, se diseña y elabora una herramienta capaz de poner a disposición c comunidad científica, y de la sociedad en general, una cantidad ingente información que es de esperar sea de provecho inmediato para varias línea: investigación lingüística, sincrónica y diacrónica: lexicografía (especialmente e que respecta a la incorporación de léxico tradicional en los diccionar sociolingüística, gramática y dialectología históricas, cronología y estratigr norma lingüística y variación, terminología, onomástica, etimología, etc. En lo respecta al ámbito, obviamente, son los investigadores en lingüística galle portuguesa los más interesados, pero no sólo, pues esta herramienta debería se indiscutible interés para los romanistas. Es importante señalar que las aplicaci no se agotan en el campo filológico, pues el manejo de datos lingüísticos sistematizados se revela un auxiliar valioso para el estudio de las migracione pueblos y de tecnología, la difusión cultural, la geografía y la historia, etc[1].

2. FUENTES

El *Tesouro do léxico patrimonial galego e portugués* está abierto a cualc tipo de material léxico que esté localizado geográficamente en el dom lingüístico del gallego o del portugués; inicialmente, se limita al portugué Portugal y al de Brasil, lo que no excluye una posterior ampliación a tod dominio lusófono. Como ya se ha avanzado, los materiales de que partimos muy heterogéneos; en los párrafos que siguen ofreceremos una descripción de tipos fundamentales.

1. Glosarios dialectales y monografías etnolingüísticas de una locali parroquia o municipio o de pequeñas áreas de mayor extensión. Una parte significativa la componen diferentes tipos de trabajos académicos, tanto doctorales y de licenciatura como investigaciones de menor entidad realiz durante el doctorado o en los últimos años de la carrera, que han sido supervis por profesores universitarios y defendidas ante un tribunal; se trata, por tanto trabajos realizados por investigadores en formación o con experiencia que si una metodología científica y están sujetos a un control por especialistas. L obras pueden ser muy distintas entre sí, no sólo en la extensión o la varied calidad de la información, sino también en el modo de disponerla y analizarla. lo que se necesitará un minucioso trabajo de homogeneización de estos mater para su traslación a la base de datos del *Tesouro*. En su gran mayoría, se trat material inédito de difícil acceso para los investigadores, pues sólo se cons alguno de los ejemplares destinados a los miembros del tribunal o depositad

[1] Véanse ejemplos, referidos sobre todo a Galicia, en Álvarez Blanco *et al.* 2006.

una biblioteca universitaria, generalmente mecanografiado o policopiado en condiciones técnicas muy deficientes; por ello, el volcado en el *Tesouro* será también un modo de facilitar el acceso de la comunidad científica a estas obras.

2. Bases de datos de los diferentes atlas lingüísticos que se ocupan de los dominios estudiados, tanto aquellos en los que están involucrados parte de los miembros de los equipos de trabajo del *Tesouro* como otros que se ocupan, total o parcialmente, de esas áreas. Es importante señalar que estas bases de datos incluyen no sólo el material que será cartografiado, sino también otro tipo de informaciones que no siempre pueden ser aprovechadas para la publicación impresa y que, así, cobrarán más visibilidad. Ciñéndonos al espacio europeo, los centros de investigación y equipos participantes disponen de los datos de las obras siguientes: los materiales gallegos y portugueses recogidos por Aníbal Otero, Armando Nobre de Guzmão y Luís F. Lindley Cintra para el *Atlas Lingüístico de la Península Ibérica* (*ALPI*) entre 1930–1936 y completados en 1947–1954; la base de datos del *Atlas Lingüístico Galego* (*ALGa*), en curso de publicación por el ILG, a partir de la encuesta realizada por Rosario Álvarez, Francisco Fernández Rei y Manuel González en 1974–1977; la *Nomenclatura de la fauna y flora marítimas de Galicia* de Carme Ríos Panisse (1977–1983) y los cuadernos gallegos destinados al nunca publicado *Atlas Lingüístico de los Marineros Peninsulares*, ambos trabajos de campo realizados por Carme Ríos Panisse y Antón Santamarina en los primeros años 70 del siglo pasado; el *Atlas linguístico do litoral português*, coordinado por Gabriela Vitorino y del que fue ya publicada la parte correspondiente a su tesis (1987); las bases de datos del *Atlas Linguístico-Etnográfico dos Açores* (*ALEAç*), en curso de elaboración en el CLUL, con datos recogidos entre 1979 y 1996; y los datos de los cuadernos y grabaciones del *Atlas Linguístico e Etnográfico de Portugal e da Galiza* (*ALEPG*), provenientes de campañas realizadas fundamentalmente por Manuela Barros Ferreira, João Saramago, Luisa Segura y Gabriela Vitorino entre 1973 y 2004.

3. Otros archivos dialectales, de diferente tipo, procedentes de encuestas sobre el terreno o por correspondencia, de temática general o restringidas a un campo semántico específico, entre los cuales cabe señalar, por su especial relevancia, el *Inquérito Linguístico* por correspondencia comenzado por Manuel de Paiva Boléo en 1942 (ILB), que se conserva en la Faculdade de Letras de la Universidade de Coimbra. El ILG cuenta también con un *Tesouro* lexicográfico en papel, en el que, entre otras aportaciones, fueron volcadas cientos de encuestas temáticas realizadas por toda Galicia por alumnos y profesores de filología románica en los años 70 y 80 del siglo pasado. De la misma forma que el ILB fue base de una importante línea de estudios dialectológicos en Portugal, el archivo del ILG fue base de algunos de los principales diccionarios gallegos modernos, a pesar de estar ambos inventarios inéditos. Nos constan otras iniciativas particulares o institucionales que dieron lugar a recogidas más o menos sistemáticas, algunas de las cuales están ya a nuestra disposición; es voluntad del equipo del proyecto localizar e integrar en la base cuantas podamos descubrir.

4. Otras contribuciones científicas que contienen léxico dialectal localización espacial, entre las cuales ocupan un papel notable diferentes artí y monografías que estudian diversos aspectos de la cultura material, en perspe onomasiológica o etnolingüística, dispersos en publicaciones que ocupan un espacial y temporal bastante amplio. Muchos de estos trabajos son obr referencia inexcusable y están relativamente accesibles y bien protegidos del o porque figuran en las bibliografías al uso; pongamos como ejemplos las pioneras de la dialectología gallega, en las primeras décadas del siglo XX, mano de los filólogos alemanes (Walter Ebeling, William C. Schroeder, K Karl Schneider, etc.) y de los entonces jóvenes investigadores del Seminar Estudos Galegos (Xaquín Lorenzo, Fermín Bouza Brey, etc.), o bie trabajos realizados por Fritz Krüger y varios discípulos por toda la Pen Ibérica. Ahora bien, hay mucha información dialectal contenida en nume trabajos publicados por autores menos conocidos y en publicaciones loc comarcales que usualmente no son consideradas en las bibliografías científic tarea del equipo su recopilación y su integración, con la clave correspondien la base de datos común.

De acuerdo con el plan de desarrollo del proyecto, en una primera f están introduciendo en la base de datos glosarios dialectales, repertorios léx monografías etnolingüísticas locales o de áreas menores (parroquias, muni comarcas, etc.). En las siguientes etapas se procesarán trabajos académic extensión breve, encuestas dialectales y artículos o capítulos de libro real desde una perspectiva onomasiológica. En último lugar se incorporará materiales provenientes de las bases de datos de proyectos de atlas lingüístic

3. ESTRUCTURA DE LA BASE DE DATOS

La base de datos que integrará todo el material se concibe con herramienta adaptable a las distintas necesidades de los usuarios, con posi de recuperar información según diferentes criterios de búsqueda e integra una aplicación para la cartografía automatizada de las respuestas.

El material léxico que se introduzca estará lematizado, por lo que el podrá realizar consultas a través de lemas (sea en gallego, sea en portugu través de cualquiera de las variantes gráficas, fonéticas o morfológicas asoc cada lema. La lematización es imprescindible en un corpus de este tipo, ya condición indispensable para poder superar la diversidad gráfica, fon morfológica del heterogéneo material estudiado. Así, una búsqueda a pa lema OVO recuperará inmediatamente no sólo la palabra *ovo*, sino las v fonéticas *ôbo, óvo, óbo* o las morfológicas *óvos, óbos, ôvos, ôbos*.

La gran heterogeneidad del material obligará a realizar di adaptaciones y reconstrucciones. Señalemos, por dar un ejemplo,

problemático de las monografías que no están dispuestas en forma de glosario, sino en texto corrido, para las que tendremos que crear un lema; así, ante un enunciado como "A pedra por trás do lar, na chaminé, /a 'bnèkra'/ [...] Num prego, na parede, está o /talêguin/ para trazerem o avio", necesitaremos crear la entrada *bonecra* bajo el lema BONECA y la entrada *taleiguinho* bajo el lema TALEIGO, a las que asociaremos la información respectiva.

La estructura nuclear de la base de datos dispondrá de los campos necesarios para recoger toda la información que se extraiga de las fuentes y también aquellos otros destinados a facilitar la búsqueda y representación de los materiales. Al final del artículo se ofrece una muestra orientativa del resultado de la búsqueda de la palabra *queixo* 'barbilla'.

3.1. Lema

El lema se corresponderá con la forma estándar, siempre que exista, y permitirá agrupar las variantes: lema (gal.) CHEMINEA / variantes {*chiminea, chaminea, cheminea...*}; lema (port.) CHAMINÉ / variantes {*chaminé, chaminei, chuminé, cheminé, cheminei...*}. Cuando no exista una forma estándar en el léxico de una lengua se acudirá a la forma estándar correspondiente en los otros inventarios: así, las variantes gallegas {*alfaza, alfacia, alfácer, alface...*} se lematizan como ALFACE, como las variantes portuguesas equivalentes. Cuando la forma estándar no exista ni en gallego ni en portugués, será necesario construir el lema a partir de las variantes recogidas y según los criterios establecidos: por ejemplo, gal. ITE 'idéntico' / variantes {*ite, itiño, itiña*}. Dentro de cada inventario, los lemas homónimos serán convenientemente diferenciados.

Algunos de los lemas pueden ser complejos (port. CIGARRO DE PALHA, gal. [*estar, vivir...*] PORTA CON PORTA) y deben ser recuperados en bloque como unidades pluriverbales, con independencia de que también se listen como sublemas bajo CIGARRO y PORTA. Estos y otros problemas y soluciones están contemplados en el protocolo metodológico al que ya se ha hecho referencia.

Consideramos necesario crear un lemario gallego, un lemario portugués y un lemario brasileño. Las formas de estos lemarios estarán vinculadas, de modo que los lemas gallegos VIÑO y PAN se relacionen automáticamente con los lemas portugueses VINHO y PÃO; de esta manera se facilita la búsqueda unitaria en la integridad del corpus y se evita la utilización de macrolemas, solución alternativa que consideramos problemática y poco operativa.

3.2. Variantes

En este campo se ofrecerán las formas léxicas tal cual aparecen en los originales, sin ningún tipo de manipulación por parte de los responsables del *Tesouro*, con la finalidad de conservar todas las variaciones fonéticas, gráficas y morfológicas contenidas en las fuentes manejadas: *queixo, queixos, quêxo, quexo,*

queijo, queixo de baixo... (variantes del lema [port.] QUEIXO 'partes do co humano').

Es probable que las búsquedas más comunes se realicen sobre este campo. todo caso, las variantes estarán vinculadas entre sí a través de los lemas (galleg portugués) y el usuario podrá obtener siempre la relación completa de todas opciones formales que fueron agrupadas bajo un mismo lema.

3.3. Transcripción fonética

En algunas de las obras que se incorporarán a la base de datos figuran, al l de las formas léxicas, trascripciones fonéticas más o menos estrictas que necesario conservar. Los alfabetos y convenciones utilizados en las transcripcic son muy variables, por lo que es preciso elaborar un sistema de corresponden que permita ofrecer a los usuarios esta rica información de forma inequívoc comprensible; para ello se usará una fuente de caracteres que no ofr dificultades de visualización.

3.4. Información gramatical

En este campo se recogerá la información sobre la clase de palabras categorías y cualquier otra indicación gramatical relevante que figure en la fu Los vocabularios dialectales acostumbran incluir anotaciones gramaticale forma de abreviatura. En las monografías descriptivas y en los atlas lingüís estas indicaciones son mucho menos comunes, pero se pueden extraer d anotaciones o, cuando existe, del contexto.

3.5. Definición

La información que se ofrece como definición es la que se encuentra en fuente, transcrita de forma literal. Las obras con formato de vocabulario mue el significado de las voces en forma de definición lexicográfica o por medio sinónimo o de la traducción en otra lengua[2]; algunas optan por la descripción por el informante, generalmente extraída de la grabación de la encuesta:

Gräniso: saraiba, pedrazo (Camariñas).

Abidueira: Abedul (O Valadouro).

Forcada: Ferramenta do labrego que ten tres gallos por baixo e un po (Silleda).

Arco de bellas: "un arco que parece qu' está con cada punta en seu reg Eso prodúcese cando chóbe è bai sol ò mismo tèmpo. È de l colores" (Guntín).

Para las obras narrativas o descriptivas, que no tienen forma de vocab indexado, en este campo se recogerá el contexto en que aparece la palab mismo ocurre cuando la palabra se registra dentro de la definición de otr entrada independiente:

[2] En muchos vocabularios gallegos anteriores a los años ochenta la definición se hace en e

Billa: Es del juego de la *billa*, es decir, de la *tala* o *toña*, consistente en hacer saltar un palito pequeño – la *billa* – y enviarlo lo más lejos posible, lanzándolo con otro más largo, la *paleta*. Se marca o traza un corro en el suelo y el otro jugador ha de impedir que la *billa* caiga dentro del corro, ya que si entra, el que lo defiende pierde y se cambian los papeles (Ancares).

Tronada: *Tromenta / Tòrmenta*. Cando a tormenta é con lóstregos dise *tronada* (Camariñas).

En los materiales que provengan de los cuestionarios de atlas lingüísticos se señalará la pregunta del cuestionario correspondiente y, si es necesario, una explicación sobre el concepto preguntado.

3.6. Ejemplos

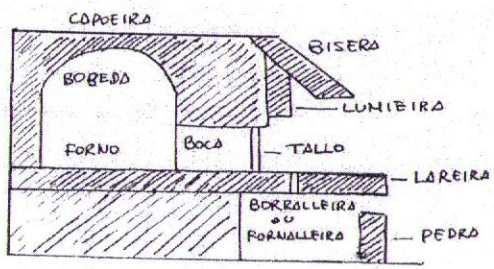
Cuando existan, se incluyen aquí los ejemplos o los fragmentos de texto con que se ilustran en las fuentes las formas indexadas. Esta información puede ser útil para conocer la adecuación de la definición o para tener información sobre el entorno sintáctico.

3.7. Refranes y expresiones fijas

Muchas de las fuentes que se utilizan para este proyecto tenían además de una finalidad lingüística un afán etnográfico. No es extraño que las definiciones se acompañen de expresiones fijas de las que forma parte la palabra tratada. Nos ha parecido conveniente recoger en la base de datos esta información, interesante también desde el punto de vista cultural.

3.8. Imágenes

Varias de las fuentes incluyen en sus páginas, especialmente en el tratamiento de determinados campos semánticos muy ligados a la cultura material, fotografías y dibujos de construcciones, herramientas, plantas, animales, etc., realizados por el propio investigador o extraídos de monografías específicas. Estas ilustraciones, en muchos casos testimonios de un mundo ya desaparecido, se incorporarán también a la base de datos, vinculándolas a las formas léxicas relacionadas y al ámbito temático correspondiente.



Ejemplo de dibujo de horno extraído de Álvarez 1974, 28.

3.9. Clasificación semántica

Todas las variantes introducidas en el *Tesouro* estarán clasific semánticamente. La adscripción a un campo semántico de las voces va a perr restringir las buscas a un determinado ámbito semántico e incluso conocer toda voces de la base asociadas a una determinada clase significativa. La clasifica semántica se realizará a partir de la propuesta por Hallig y Wartburg par francés[3] y aprovechando las agrupaciones semánticas que aparecen en alguna las obras. En principio se ha comenzado a trabajar con dos niveles jerárquico agrupación (*aixola*: oficios > carpintero).

3.10. Geosinónimos

Los geosinónimos son las formas léxicas asociadas a distintos lemas cc mismo significado: QUEIXO, CARRINHO, BARBA, MENTO, QUEIXADA, QUEI PAPO, BARBELA, BARBADELA, PERA, CARRO, QUEIXELO, BARBAROTE, BICO.. conocer los geosinónimos vinculados con una determinada forma sabremos distintas denominaciones que se recogen en la base para un mismo concepto vinculación de los geosinónimos se realiza a través de un macrolema que sirve relacionar todos los lemas, con sus variantes, que remiten a un mismo concepto

3.11. Fuente

Cada una de las fuentes utilizadas se identificará con un código que perm conocer el origen de las variantes. De esta manera se pueden recuperar los c completos de la fuente: autor, título, tipo de obra, fecha, lugar o área a la corresponden los datos, etc.

3.12. Identificador geográfico

Las fuentes materiales que se incorporan a la base están identific geográficamente. Este identificador es el que permite conocer los dominios er se registran las variantes y también representar la distribución de las fo (variantes y lemas) en un mapa. Para que la representación de los datos sea vi se ha pensado en utilizar como base los mapas de municipios para el terri europeo y las mesorregiones para Brasil. El usuario puede conocer primeran en forma de texto las localizaciones exactas de las variantes y lemas; si q conocer con mayor precisión las localidades en que fueron recogidos los c deberá acudir a la ficha bibliográfica de la fuente.

4. CARTOGRAFÍA

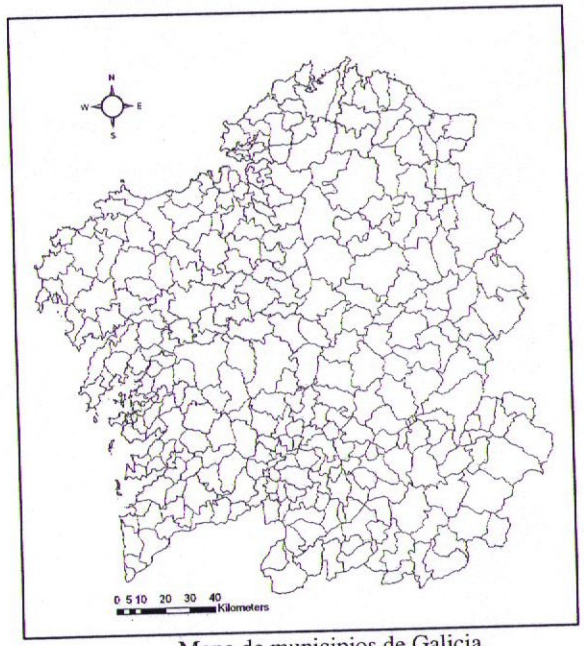
La información sobre la localización geográfica de las formas se ofrece la ficha de resultado de la búsqueda como texto y también con una represent geográfica.

[3] Hallig/Wartburg 1952,1–94.

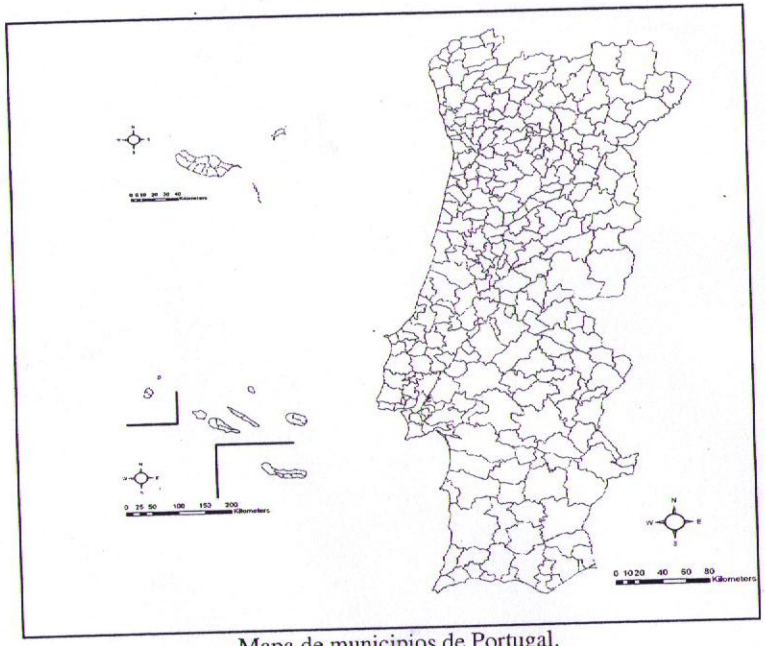
La aplicación ofrecerá la posibilidad de que el usuario pueda ver la distribución geográfica de las formas sobre un mapa. Como se ha indicado más arriba, se ha escogido en principio una representación basada en áreas y no en puntos; de este modo, sobre el mapa aparecerán destacadas con color aquellas áreas (municipios o mesorregiones) en las que se registró la forma o formas buscadas. En el dominio lingüístico gallego, la red de unidades mínimas está formada por los 315 municipios (*concellos*) de Galicia, a los que se suman los municipios gallegohablantes de las provincias fronterizas (Asturias, León y Zamora). La red portuguesa está formada por los 277 *concelhos* del Portugal continental y las 11 islas para los archipiélagos de Azores y Madeira. En Brasil, la red está formada por 142 mesorregiones de tamaño variable, distribuidas en los 26 estados.

Así, quien consulte la base podrá hacerse una idea sobre la extensión del uso de la voz y también sobre la continuidad de su distribución. Al pasar el puntero sobre estas áreas sabrá también en cuántas y cuáles obras se ha registrado y con qué variantes. Con seguridad este complemento cartográfico será de gran ayuda para estudios geolingüísticos e históricos.

Mapa de mesorregiones de Brasil.



Mapa de municipios de Galicia.



Mapa de municipios de Portugal.

5. PARTICIPANTES

En este proyecto participan investigadores gallegos, portugueses y brasileños. La coordinación general del proyecto se realiza desde el Instituto da Lingua Galega (ILG) de la Universidade de Santiago de Compostela bajo la dirección de la profesora Rosario Álvarez. La sección portuguesa de este proyecto está integrada por investigadores pertenecientes al Centro de Linguística da Universidade de Lisboa (CLUL) y al Centro de Linguística Geral e Aplicada da Universidade de Coimbra (CELGA), capitaneados, respectivamente, por João Saramago y Clarinda Maia. En esta ocasión, nos centraremos en las ramas europeas del *Tesouro do léxico patrimonial galego e portugués*, pues la red brasileña, coordinada por Sílvia Brandão, de la Universidade Federal de Rio de Janeiro, y Vanderci Aguilera, de la Universidade Estadual de Londrina, todavía está en proceso de constitución.

5.1. Equipo gallego

La rama gallega se ocupa, como ya se ha dicho, de la coordinación general del proyecto, lo que implica dirigir el diseño informático, el modelo lingüístico y su metodología, la representación cartográfica y la presentación vía Internet de los resultados; en todas estas cuestiones participan, de una forma u otra, todos los miembros del equipo internacional. En este aspecto, las tareas inmediatas son las siguientes:

a) Diseño y gestión de la base de datos a la que volcar la información contenida en las fuentes; diseño de la aplicación para la introducción y revisión de datos en línea; diseño de la aplicación que permita la representación en un entorno geográfico de los datos léxicos; diseño de la herramienta web de consulta de datos (consultas textuales y representación cartográfica).

b) Elaboración de un protocolo metodológico, detallado y unificado, de actuación homogénea para todos los lingüistas miembros del equipo (tratamiento de los datos e inclusión en la base, criterios de lematización, criterios de clasificación semántica, etc.), publicable como manual para proyectos semejantes.

c) Diseño y elaboración de un prototipo novedoso de diccionario dialectal electrónico, con cartografía, exportable. El prototipo se acompaña de un sistema de búsquedas completo.

El comité gallego, correspondiente de los comités portugués y brasileño, está coordinado por Xulio C. Sousa. Lo integran, además de la coordinadora general (Rosario Álvarez), María Álvarez de la Granja, Francisco Fernández Rei y Ernesto X. González Seoane (profesores de la Universidade de Santiago), Xandra Santos Palmou (Universidad de Salamanca), Xosé Afonso Álvarez Pérez (actualmente en

el CLUL, Universidade de Lisboa) y dos investigadoras en formación (Negro Romero y Mónica Martínez Baleirón). Las tareas inmediatas del c gallego son las siguientes:

a) Inventario y reunión de todas las fuentes que contienen léxico rec con metodología científica y referencia geográfica. Elaboración y publicac internet de la relación bibliográfica, con la información pertinente p. identificación y localización de cada fuente.

b) Copia automática (OCR) o manual de los textos seleccionados y tr de la información que contienen a la base de datos. En el primer trieni previsto el volcado de tesis doctorales, memorias de licenciatura y tr análogos, que suman aproximadamente un centenar, distribuidas por tc territorio gallegófono.

5.2. Equipo portugués

El comité portugués está coordinado por João Saramago (Universid: Lisboa, CLUL) y por Clarinda Maia (Universidade de Coimbra, CELGA) y forman parte las profesoras Isabel Santos, Maria José Carvalho (CEL Gabriela Vitorino (CLUL). Además, se integran en el equipo dos doctor Carmen Gouveia (CELGA) y Sandra Pereira (CLUL). Por último, también parte Xosé Afonso Álvarez Pérez (CLUL), que, como ya se ha referido, es ta miembro del comité gallego.

La primera fase del proyecto en su vertiente portuguesa será la digitali: reconocimiento óptico de caracteres y etiquetado de 122 tesis doctorale: licenciatura conservadas en las universidades de Coimbra y I correspondientes en su gran mayoría a investigaciones de campo realizada los años 40 y 70. La red es amplia y cubre las principales áreas del P continental e insular. Aunque normalmente el objeto de la tesis es la desc lingüística de una localidad o de un pequeño grupo, algunas de ellas estuc tema específico a lo largo del país[4].

En definitiva, deseamos que este proyecto que ahora emprende sus pr pasos ponga a disposición de científicos y ciudadanos un amenazado patr cultural que, aunque no sea tangible como las piedras de las catedrale: pinturas de los artistas, tiene una importancia de primera magnitud para n comunidades y constituye una riquísima fuente para estudios lingüíst etnográficos de todo tipo.

[4] Por ejemplo, los trabajos académicos de Alberto Pires (1959) *A "Cegonha" – com para o estudo linguístico-etnográfico dos aparelhos de elevar água*, de M. da Conceição Mot mismo año, *O linho-cultura e fabrico (Estudo linguístico etnográfico-folclórico)* o de M. F. F Tavares (1973) *A terminologia do moinho.*

Lemas: *queixo* Bras., Gal., Port.
Clasificación semántica: corpo humano > cabeza; alimentos
Variantes: *queixo* (gal.), *queixelo* (gal.), *queixo* (port.), *queijo* (port.), *quêxo* (port.)
Fuentes: **Bras.**; **Gal.**: Todas; **Port.**: Todas
Gal. queixo {*queixo*} COD.: barbilla, mentón, p. 59 GOI.: mentón; p. 31. MEL.: mentón; p. 33. NOV.: punta da queixada de baixo; p. 62. RAM.: debaixo da boca, no queixal, está o QUEIXO, que o parte en dúas mitades; p. 60. TOB.: la barbilla; p. 46. ALPI s.m. Cuest. 591 'barbilla' [parte saínte da mandíbula inferior das persoas] {103, 104, 110, +}. ALGA s. Cuest. 2115 'queixo' [parte saínte da mandíbula inferior das persoas] {Sada, A Coruña, Malpica de Bergantiños, A Laracha, Ponteceso, Laxe, Carballo, Carral, +}. **Port** queixo {*queixo, queijo,quêxo*} GERM.: [kˈɐjʒu]; p. 355. ESC.: [kˈeʃ]; p. 335. GLO.: [kˈeʃ]; p.142. M. VELHA: queixo; p. 222. MT. GORDO: quêxo; p. 280. OLH.: [kˈeʃ]; p. 176. ALEPG Quest. 429 'queixo' [nas pessoas, parte saliente por baixo do lábio inferior] {Mértola, Alcáçovas, Silves, Peniche, Alenquer, Campo Maior, Sines, Mação, Arganil, Sátão, Castelo de Paiva, Alfândega da Fé, Montalegre, Melgaço, Madeira, [+]}. ALEAÇ Quest. 429 'queixo' [nas pessoas, parte saliente por baixo do lábio inferior] {Faial, Graciosa, São Miguel, Santa Maria, [+]}. ALPI s. Quest. 591 'barbilla' [nas pessoas, parte saliente por baixo do lábio inferior] {Melgaço, Paredes de Coura, Caminha, Viana do Castelo, Ponte da Barca, Vieira do Minho, [+]}.
Geosinónimos: **Gal.** balbarote, barba, barbarote, barbirolle, bico, cacín, carriño, carrillo, [+]; **Port**: barba, barbadela, barbela, carril, carrilho, carrinho, papo, queixo, [+].
Distribución geolingüística: Bras., Gal., Port.

Muestra de ficha de resultados de la búsqueda de *queixo* ('barbilla')

REFERENCIAS BIBLIOGRÁFICAS

Álvarez Blanco 1974 = Rosario Álvarez Blanco, *Notas lingüísticas e etnográfic Ramirás*, Tesis de licenciatura (inédita), Univ. de Santiag Compostela.

Álvarez *et. al.* 2006 = Rosario Álvarez *et al.* (eds.), *Lingua e Territorio*, Santiag Compostela, Consello da Cultura Galega, pp. 257–308.

Hallig/Wartburg 1952 = Rudolf Hallig y Walther von Wartburg, *Begriffssystem Grundlage für die Lexikographie. Versuch eines Ordnungssch* Berlin, Akademie-Verlag.

PRESENTATION OF A DIGITAL CORPUS OF TRADITIONAL LEXICON *TESOURO DO LÉXICO PATRIMONIAL GALEGO E PORTUGUÉS*

ABSTRACT

The *Tesouro do léxico patrimonial galego e portugués* is an international project coord by the Instituto da Lingua Galega (Universidade de Santiago de Compostela). It aims to colle make available to the scientific community different lexicographical materials, with geogra reference, from Brazil, Galicia and Portugal, which represent an important part of the herit these countries, often threatened by the abandonment of traditional ways of life. This online da will include both published and unpublished material of different type (dialect glossaries, ethnolin monographs, databases of linguistic atlas...); it will allow queries by headword or variant and also incorporate an application for automatic mapping of the results selected by the user.

LUCRĂRI RECENT APĂRUTE LA EDITURA ACADEMIEI ROMÂNE

INSTITUTUL DE LINGVISTICĂ „IORGU IORDAN – AL. ROSETTI", **Dicționarul limbii române (DLR)**, serie nouă, tomul I, partea a V-a, litera **D (DEȚINUT – DISCOPOTIRIU)**, 2007, 386 p.

INSTITUTUL DE LINGVISTICĂ „IORGU IORDAN – AL. ROSETTI", **Atlasul lingvistic român pe regiuni. Muntenia și Dobrogea**, vol. V, coord. Teofil Teaha, 2007, XX p. + 116 hărți + 135 planșe.

INSTITUTUL DE LINGVISTICĂ „IORGU IORDAN – AL. ROSETTI", **Limba română, limbă romanică. Omagiu acad. Marius Sala la împlinirea a 75 de ani**, coord. SANDA REINHEIMER RÎPEANU și IOANA VINTILĂ-RĂDULESCU, 2007, 614 p.

INSTITUTUL DE LINGVISTICĂ „IORGU IORDAN – AL. ROSETTI", **Formarea cuvintelor în limba română din secolele al XVI-lea – al XVIII-lea**, coord. MAGDALENA POPESCU-MARIN, 2007, 412 p.

INSTITUTUL DE LINGVISTICĂ „IORGU IORDAN – AL. ROSETTI", **Sintaxa limbii române în secolele al XVI-lea – al XVIII-lea**, coord. MIOARA AVRAM, 2007, 328 p.

GR. BRÂNCUȘ, **Studii de istorie a limbii române**, 2007, 284 p.

MIOARA AVRAM, **Studii de sintaxă a limbii române**, 2007, 194 p.

NICOLAE SARAMANDU, **Aromâna vorbită în Dobrogea. Texte dialectale. Glosar**, 2007, 660 p.

DANA-MIHAELA ZAMFIR, **Morfologia verbului în dacoromâna veche** (secolele al XVI-lea – al XVII-lea), partea a doua, 2007, 380 p.

SANDA GOLOPENȚIA, **Limba descântecelor românești**, 2007, 238 p. („Limba română" 8).

MARIA ILIESCU, **Româna din perspectivă romanică**, 2007, 410 p.

DUMITRU LOȘONȚI, **Certitudini și ipoteze etimologice**, 2007, 180 p. („Etymologica" 26).

IULIA MĂRGĂRIT, **Comentarii etimologice și semantice**, 2007, 396 p. („Etymologica" 27).

ISSN 0071–6855