VIII. New developments in the lexicography of individual languages since 1990 II: The Romance languages

39. Iberoromance I: Historical and etymological lexicography

1. Introduction
2. Individual languages: An overview
3. Individual dictionary types
4. Conclusion: Innovation, prospects
5. Selected bibliography

1. Introduction

"Quiconque s'intéresse à l'étymologie et à la chronologie des mots ibéroromans, se trouve dans une situation fort compliquée et peu satisfaite" wrote Dieter Messner (DCE, 1976, v) thirty seven years ago, and more recently P. Alvarez de Miranda (2003: 65, with reference to Menéndez Pidal, "El diccionario que deseamos observa que, con respecto a Spanish de nuestro diccionario que se desea ha en 1945 sigamos descansando 57 años después." Whether or not a view of achievements in the field of the historical and etymological lexiconography (HEL) of the Iberoromance languages must be met with the same glee and doom today, is a question we will endeavour to answer in this contribution.

What has already been said on the history of dictionaries of individual languages in the survey articles in HSK 5.2 (1990): Haenssle a and b, Pensado, Woll, in LRL Alvar Ezquerra (1992), Brea (1994), Rico Soló (1991), Verdelho (1994); and in HSK 23.1 (2003): Bolle [on creoles], Colón, Monjou, and Müller is taken as known; projects completed before 1990 will not be mentioned again.

The practical problem of presenting several languages and numerous types of dictionaries simultaneously in one article is solved as follows: Chapter 2 contains a brief overview of the languages and projects covered, with forward references to Chapter 3, which presents the individual projects, organized by dictionary type.


2. Individual languages: An overview

As 'Iberoromance', we deal in what follows with, on the one hand, Galician (G), Portuguese (P), and Galician-Portuguese；for the hard-to-draw line between them (cf. HGP 885f.), Asturian-Leonese (AL), Spanish/ Castilian (S), Navarrese-Aragonese (NA), Catalan – in the broader sense: Catalan, Valencian, Balearic – (CBV) as languages of the Iberian Peninsula; on the other hand, under P and S, naturally the worldwide varieties of these languages also are included. We thus adhere grosso modo to the practice of LRL (cf. Metzeltz-Winkelman 1992, Osenkop/ Winkelman 2003). In the articles in Chapter 3, this sequence of languages is also maintained in the discussion of individual dictionaries.

The individual languages and projects are as follows:

- P: Araujo 2001: 3.3; CIPM: 3.3; DDP: 3.10, DLEP: 3.2, DOELP: 3.11, DVPM: 3.3, Houassi 2001: 3.12, IPVM: 3.9, VHEP: 3.5, VPM: 3.3.
- S: Alvar Ezquerra 1997: 3.5, DCECH: 3.2, DEDA: 3.7, DELAM: 3.6, DEM: 3.3, DEM: 3.9, DeonHisp: 3.11, DESE: 3.6, DETEMA: 3.6, DHA or Gledjen 1997: 3.4, DHECRI: 3.4, [not consulted], DHEHCAN: 3.4, DHLE: 3.1, DPCA: 3.7, Kasten/Cody: 3.3, LHMEX: 3.4, LHP: 3.5, LHLAM: 3.7, NDLHE: 3.1, NTLE and NTMLE: 3.10, Pezzi 1995: 3.5, Rosal 1613ca: 3.10, Siebenjünger 1993: 3.5, TLECan: 3.10, TLEMAR: 3.6, TLEPR: 3.10, Traisa y Antovezca 1997: 3.6, TVIL: 3.6.
- CVB: DVCB: 3.1, DVCBI: 3.1, DECat: 3.2, DHIVAM: 3.1, DHVC: 3.6, DMLC: 3.5, Gracia Giner 1994-98 [not consulted], OnomCat: 3.11, Vocabulari 1502: 3.10.
- P-based Creoles: DECPA: 3.2.

3. Individual dictionary types

3.1. Comprehensive historical dictionaries

In this chapter, "Comprehensive" is used to mean: aiming to cover the vocabulary in its entirety. No such dictionary exists for Galician and Portuguese; for E it is available in rudimentary form in Diccionario histórico de la lengua española (DHLE): since 2006, however, the RAE Nuevo diccionario histórico de la lengua española has been under development: in the initial 2006 - 2008 phase of the project, a smaller corpus is to be extracted from the CORDE corpus, lemmatized and supplemented with entries from modern dictionaries in order to form the basic material; in parallel, a plan for about 50,000 lemmata and the microstructure will be created. In 2018, a first version of the dictionary is supposed to be available, cf. Pasqual Rodrigues / García Pérez 2007, 2008, Pasqual 2009.

DVCB can still refer to the Alcover/Moll Diccionario catala-valencià-balear (DVCB) ("una de las más admirables y elegantes producciones de la filología románica", Colón 2003: 371); it was designed as comprehensive, however, only in terms of vocabulary and not of the documentation of, for example, each century. The second edition appeared in 1993, and since 2001 - 2002 it has been accessible on the Internet (DVCB); articles accessible singly, it cannot be "browsed". For DHIVAM, cf. 3.9.

3.2. Comprehensive etymological dictionaries

There is no etymological dictionary for Galician (Bre 1994: 127); for P, P.J.P. Machado's Diccionario etimológico do vocábulo português (DELP) is available (the 8th edition appeared in 2003); in Asturian, the Diccionario Etimológico da Língua Asturiana (DELIA, Garcia Arias) is being compiled; for E, the Coromías/Pascual Diccionario etimológico castellano e hispánico (DCECH, completed in 1991) comes closest to filling this role; cf. (Agustí 2000, Baldring 1998): this etymological dictionary is arranged by lexical family, and it is enormously rich in content. Causes of criticism have been the neglect of semantics, derivatives (particularly in comparison to FEW; the DCECH was not intended to be such a dictionary, cf. foreword) and dialectology, and the lack of distinction between documentation and commentary; the non-objective treatment of other researchers' opinions is unfortunate. For CVB, much the same applies to Diccionario etimológico complementario de la lengua catalana (DECat, completed in 1991, supplement and index added in 2001; cf. Colón 2003: 372f., Rico / Solà 1995: 149): it represents an improvement to the coverage of dialects; the polesmics are worse.

One project that, as a "new REW", also covers Iberoromance languages is Diccionario Etimológico Román (DÉRoma) Buchs/Schweickard, presented in Innsbruck, September 2007; the first 70 articles have been published on the Internet since 2008, cf. http://www.atll.irdormer.org/). In the field of creoles, five Portuguese-based creole dictionaries compared etymologically in J-P. Rouge's Diccionario etimológico dos creoles portugueses d'Africa (DECPA, 2004).

3.3. Historical and etymological dictionaries (HEDs) of defined periods

On a contemplated Vocabulário histórico-etimológico do português medieval, cf. Araujo (2001: 52f.) and 3.9 (IVPM). Another project on P Diccionario de verbos do português medieval (DVPM), is, for now, a Web-published dictionary of verbs only, but was conceived as being part of a future comprehensive dictionary of medieval P. It is based on the CIPM electronic corpus of the University of Lisbon.

B. Müller's Diccionario del español medievo (DEM, 1987) is a highly ambitious
work, both philological and also etymological and inter-Romanially comparative. Its aim is complete coverage of Old Spanish up to 1400 (including Old Aragonese), by 2007 just three volumes had been published, with a range (A-alinhacat) corresponding, on the basis of DRAE, to only about five per cent of the alphabet (on DEM, cf. Dворкин 1994: 411f, Müller 2003: 392, on DEM cf. 3.9). The Heldenberg project has been suspended at the moment, material is being transferred to La Laguna, Tenerife, for further processing.

In contrast, the Tentative Dictionary of Modern Spanish (VHEEP) is the first single-volume dictionary of Old Spanish that evaluates 86 sources and previous glossaries. More than a few of the lemmata taken from, for example, Oesterhödt, ought to be classified as Medieval Latin.

3.4. HEDs of defined geographical areas

The draft of a Diccionario Histórico Americanos (1492–1836) (DHA or Gilegren 1997) does not describe a project to be implemented in the near future, but is rather a scenario for a conceivable HED on the Americas (cf. also Lüdtke 2005). Concerning individual countries, M. Morena’s Diccionario histórico-etimológico del habla canaria (DHEECan) 2001) studies the various types of canarisms diachronically giving examples and etymology. On Diccionario histórico del español de Canarias (Corrales/Corrubella 2001) and P. Quesada Parra’s Diccionario histórico del español de Costa Rica (1995), cf. Álvarez de Miranda (2003: 67). The Lexico histórico del español de México by Company/Melis (LHEMex, 2002), without definitions and datings) the contexts and syntactic usage of the lemmata with selected examples, based on Documentos lingüísticos de la nueva España (C. Company, México, 1994); it focuses on the colonial period. On TLECan and TLEPRI cf. 3.10.

3.5. HEDs of contact languages; Stratified dictionaries

A. G. da Cunha’s Os estrangeirismos da língua portuguesa: vocabulário histórico-etimológico (1998) is the first of a projected version of a planned Diccionario histórico-etimológico dos estrangeirismos da língua portuguesa (DHEEPP), which will try to collect all borrowings to or from Portuguese, with extensive data on the history of the words. Work has been or is being done on individual contact languages and data, e.g.

1) Medieval Latin: Léxico hispánico primitivo (LHP, 2003) stems from a Menéndez Pidal project: it seeks to document precursors of Romance lexemes and adoptions into the Medieval Latin (and Arabic) of northern Spain, principally in Asturias-León, from the eighth to the twelfth century; the corresponding dictionary of the same period for the Glossário de vocábulos e expressões do lusitâneo (GLOML, 1960–85; A–D, cf. Rico/Soló 1995; 151) for the ninth, tenth and eleventh centuries, supplemented in 2001 (Kasten/Cody 2001) is in the area.

2) Arabic: the definitive works that have been published on the topic are, firstly, R. Kesler’s Kleines vergleichendes Wörterbuch der Arabismen im Itersromanischen und Italienischen (1994), which takes into consideration arabics each from P, S, CVB and Italian and compares them with the results in the other languages. The introduction alone is noteworthy: this is by no means a "small" work; secondly, F. Corriente’s Diccionario de arabismos y voces afines in Iberoromanes (1999, 2003, "obra de gran rigor científico de elaboración grandiosa, gran especialista", Haensch/Ormaechea 2004: 145, cf. Kesler 2002), is a comprehensive etymological dictionary that covers all Iberoromance idioms; finally, Peri 2005 can also be consulted as a supplement to the actual dictionaries.

3) French: on the project for a Diccionario de Galicismos de Costa Rica (Kasten/Cody 2003).

4) American languages: Nativism in general are handled in the Vocabulario de indigenismos en las Crónicas de Indias (Alvar Ezquerra, 1997); they are taken from 22 crónicas and presented in contexts, finally, especially for Quechua, Siebenhügener (1993) has the character of a dictionary.

3.6. HEDs on individual varieties; technical and special purpose HEDs

Diccionario español de textos médicos antiguos (DETEMA, 1996), compiled under the direction of M.T. Díez and exhaustively evaluates 32 primary sources on medicine with texts mainly from the thirteenth century. L. Nieto Jiménez’s Tesoro lexicográfico del españo (ESTLE, 2002) uses 12 texts from the years 1538–1722 dealing with mariners’ technical language. Since these are largely of a lexicographical nature, it could also be included in 3.10. It is meant to be seen as a preliminary work to NTLE. M. Chamarro’s Tesoro de vilanismo (DEVA, 2002) is a historical (partly etymological) dictionary recording the slang of the "clases marginales" in the sixteenth and seventeenth centuries and documents a vocabulary which is an encyclopaedic, rather than strictly etymological dictionary. Under a lemmat as Caracollo Código Nero, for example, it gives legal texts at length rather than, for example, earliest dates or quotations.

A great advance in Valencian HEL is marked by J. Martí Mestre’s Diccionario histórico del valenciano colloquial (DHVC, 2006), which for one stage (seventeenth to nineteenth century), documents a (diaphasic) variety of information from a large number of primary sources. Like TLEMar, it is to be seen as a preliminary work to a more comprehensive etymological dictionary.

3.7. Dictionaries of particular types of text or individual authors


3.8. Dictionaries of particular parts of speech

For DVP (P), cf. 3.3. For S, D.A. Pharies’s Diccionario etimológico de los sufijos españo (DESE, 2002), which principally discusses the etymologies of derivation suffixes and the derivation of words, containing elements all stages of word history without claiming to be a complete historical dictionary, and A.M. Serradilla Castaño’s Diccionario sintáctico del español medieval: verbos de entendimiento y lengua (1996), a small dictionary on verb construction in Old Spanish that also provides examples, has been published.

3.9. Microstructurally abridged dictionaries

A. G. da Cunha’s Índice do vocabulário de Portugal (1990, 1992) is a complete, abridged, preliminary version of a planned dictionary (cf. 3.3); it covers texts from the thirteenth to the fourteenth century, with headwords in the original characters and entries containing only dating and text reference points. The CD-ROM Vocabu- lário do português medieval (VPM), with a total of 500,000 lexemes, has been published from the Fundação Casa de Rui Barbosa since 2006; consultation and search facilities are outstanding and contexts are given for each form. VPM can therefore, be consid- ered complete. DEM: Tesoro DEM Infor- matizado (Tesoro DEMI) is a preliminary on-line publication of the planned lemmata of DEM (cf. 3.3) with text references and first dates; the first lemmata to be accessible in September 2007 where the ones beginning with B, D, F, and G. On Endize de bocables de l’aragonés (Nagore Lain 1999), cf. 3.10.

Dictionaries of the idioma valenciano modern (DHIVAM, R. García Moya, 2006) is a propagandistic work. Directed specifically against Catalan language policy, it seeks to document the linguistic autonomy of Valencia. Texts ranging from medieval to twentieth-century, the author selects a large number of lemmata to produce a list of words with the appropriate evidence, and date, but without a definition. Among them he inserts lexemes (e.g., catá- lano), this time with a definition but without historical data.

3.10. Older dictionaries; The ‘dictionary of dictionaries’ genre

Dictionaries of past centuries are an indispensable source for present-day research (cf. the practice of Französisches etymologisches Wörterbuch or Lessico etimologico italiano; conversely, failure to consider works of the early modern era (e.g., for example, DCVB, has been criticized, cf. Rico/Soló 1995: 147). To- day many older works are made more easily accessible by being included in corpora such as CORDE (e.g., Coordinamento), or produced in facsimile or in print (e.g., reprint of Francisco del Rosal ca. 1613, S, of the Vocabulari 1502, CVB) (cf. Haensch/Ormaechea 2004: 142).
One particular way of presenting—and above all of reworking—such sources is the 'dictionary of dictionariers' genre; the idea goes back to the pioneering work of Gili Gaya (1947–60). In G and GP, Diccionario de dicionariares (DDG, 2000, 2003) and Dicionario de dicionarios do galego medio (DDGM, 2006) have been published. In P, the most comprehensive work of this genre, D. Messner’s Dicionario dos dicionariares (DOELLP, 1994–96); to date 21 vols; A. H. K. L. N. O. Q. U. is in progress, in S, dicionarios of individual countries, on the one hand, have been published; Tesoro léxico de España y de las américas by C. Corrales Zumárraga et al. (TLECAN, 1992) has about 200 mainly dialectological sources; Tesoro lexicográfico del Español de Puerto Rico by Morales/Vaquero (TLEPR, 2005), is based on 59 sources, including data from a language atlas. Its approach is not contrastive, i.e. lexemes also occurring in other Spanish-speaking countries are included. On the other hand, there is Nuevo tesoro lexicográfico de la lengua española (NTELLE, cf. Nieto Jiménez 1992), which draws from 66 dictionaries from 1726 to 1992. All these dictionaries (and others also available on the Internet), not to be confused with NTELLE: Nuevo tesoro lexicográfico del español (Nieto Jiménez/Aral Ezquerro), which documents all Spanish lexemes occurring in glossaries and dictionaries from the 14th century to 1726 (11 vols, 2007).

Preliminary work on a planned Tesoro del léxico gallego-asturiano, compiled by Endize de bocales de l’aragonés (Nagore Lain, 2004; 4 vols) and Fuens lexicograficas de l’aragonés (Nagore Lain, 1998), which can be considered as a provisional annotated bibliography to it. It maintains lists of diconarios (from 1641) and collections of dialect vocabulary. 3.11. Onomastico and deonomatico HEDs Here, there are new creations for nearly all idioms: Boulon Agreló’s Anticontrastive metalinguistic galega (1999) is of a dictionary nature in its main section; it evaluates seven collections of documents from the eighth to the twelfth century. J. P. Machado’s Dicionário onomástico e etimológico da língua portuguesa (DOELLP, 1984) has had two new editions, 1993 and 2003. In AL, La onomástica asturiana baiomediaval, (Viejo Fernández 1998) documents Asturian personal names from the thirteenth to the fifteenth century and examines their formation. The main section (from p. 269) is of a dictionary nature. For progress and desiderata, see Corominas (2003: 393) and Celdrán 2003. M. Belasko’s Dicionario etimologico de los nombres de los pueblos, villas y ciudades de Navarra (1999) explains and compares the place names of Navarra and provides early evidence about them. Onomasticario Catalán: els noms de lloc i noms de persona de totes les terres de llengua catalana by Corominas/Mascaro Passarrius (OnomCat, 8 vols, 1989–97) has been completed; vast in and of itself, it simultaneously marks the peak of Joan Corominas’ immense production.

Gallarín/Galarrín present the first forays of deonomastic research into S with Deonomástica hispánica (DeonHisp, 1997) a partly monographic, partly lexicographical, occasionally historical work (cf. Lengert 2000 for very harsh criticism). Reinhardt is planning a work on Iberoromance deonomastics that also encompasses dichrony (project description published in 2010). 3.12. Synchronic dictionaries with historical or etymological components The publication of Dicionario Houaiss da língua portuguesa (Houaiss 2001, 2007) marked a small revolution in Portuguese—no Iberoromance—a synchronic lexicography: for an enormous amount of lemmata, it gives etymological information, dates, in part also older forms, using data from, among other works, IVPM (cf. 3.3), and this publication was praised enthusiastically by Messner (2002); Barreto (2006) also expresses his respect, but at the same time criticizes errors in etymology, dating and classification as brasilistas, errors that might have been avoided by referring to currently available literature (e.g., DCECH). This in no way detracts from the importance of the dictionary, but shows once more that in historical lexicography caution must be exercised when using data on word history from synchronic dictionaries.

4. Conclusion: Innovation, prospects It must be stated that, in recent years, a substantial amount has been achieved for all the language groups discussed here, but with a certain quantitative leap in importance for the Spanish language. Four trends can be discerned: 1) the compilation of large and very large dicionarios (‘obras gigantescas’). The most ambitious one in progress is certainly DEM; in the project stage, NDHLE. The boom of something seminal in etymology and word history being achieved here is counterbalanced by the disadvantage of the very long publication cycle; 2) the compilation of restricted HEDs on single periods, regions, authors or text types. Müller (2003: 392) inveighs against these wrongly, in our opinion; they offer the benefit that, within a few years, alphabetically complete works of reference can be made available to students and researchers, also from other fields. As desiderata in this area, a dictionary of the Spanish of Siglo de Oro (Haensch/Schmitt 2004) 152 and a dictionary of Old Catalan (Haensch 1990b: 1777, cf. Colon 2003: 374) have been mentioned; 3) work on ‘dicionarios de dicionarios’ (or Tesoros lexicográficos, cf. 3.10) is noticeably dynamic in the Iberoromance field. For Catalan (in the narrower sense) there is a desideratum here; 4) the decisive, most radical innovation of the last few years has been that dictionaries, on the one hand, are being published electronically on the Internet and electronic corpora (such as ADMYTE, CORDECH, TMLIG) or philologically high-level publications, on the other hand, are being linked to dictionary projects, cf. for example, the CIPM-DVPM tandem (analogous to Italian: OVL database-TLIO dictionary) or the future CORDE-NHLE. The latter way to develop HEDs is perhaps the greatest recent innovation and the possibilities are far from being exhausted.

5. Selected bibliography


VIII. The Romance languages


40. Iberorománica II: Synchronic lexicography

1. Introduction
2. General monolingual dictionaries
3. Some specialized monolingual works
4. Bi- and multilingual dictionaries
5. Portuguese dictionaries
6. Catalan dictionaries
7. Summary
8. Selected bibliography

Introduction

The present article focuses on a brief account of the recent development, the present state and future perspectives of the synchronic lexicography of the principal Iberorománica languages. Relevant overviews include Geldi (2003), Haensch/Omellaha (2004) and Anglada (2005). The publication year of HSK 5.3 (1991) coincides almost exactly with the technological revolution and widespread usage of the World Wide Web, hyper-text, and other new media. Methodological and theoretical innovations include digitalization of the bases of dictionaries (electronic corpora, e.g. the Corpus de Referencia del Español Actual (CREA) of the Spanish Academy) and of the dictionaries themselves, as well as the general inclusion of taboo words, improvement of definitions and of the treatment of syntagmatic information. In the early 1990s, the large-scale production of electronic dictionaries began. Before that date, there were few dictionaries on diskette. Afterwards, three new forms of dictionaries appeared: the electronic dictionary, the CD-ROM dictionary, and the online dictionary. These new types can be designated as digital dictionaries as opposed to the traditional analog ones. Digital and analog dictionaries differ in various respects. Both have their advantages and disadvantages. Digital dictionaries offer easy, fast and cheap access to information and have no space limitations, but they require the necessary equipment and skill. The advantages of digital dictionaries include the possibility of electronic processing, extended search functions, easy selection, recuperation, reproduction, transport, and interactivity. The disadvantages lie in their lack of reliability, usability and stability (Geldi 2003: 314–315). Cf. Chuchuy/Moreno (2002) for a critical review of four Spanish CD-ROM dictionaries of the 1990s (DRAE 1995, DUE 1996, GDLE 1996, Vox 1997).

Different structures must be distinguished (Geldi 2003: 321–326). The hyperstructure comprises the overall organization of information in both types of dictionaries. This organization may be accessible via menus, but many digital dictionaries lack the traditional presentations such as foreword, introduction and explanatory notes or usage guides. Other internet dictionaries form part of larger portals. As for the macrostructure, digital dictionaries are in principle subject to constant alteration and extension, while analog dictionaries constitute closed sets of entries. On the other hand, many internet dictionaries offer only the requested article and not the complete data: in many cases, one cannot leaf through the digital dictionary as through paper versions. Microstructures are quite similar in traditional and digital dictionaries.


Jan Reinhart, Tübingen (Germany)
Dictionaries
An International Encyclopedia of Lexicography

Supplementary Volume:
Recent Developments with Focus on
Electronic and Computational
Lexicography

Edited by
Rufus H. Gouws
Ulrich Heid
Wolfgang Schweickard
Herbert Ernst Wiegand