

Section 6 : Lexicographie

ALDEA, Maria : DES EMPRUNTS LEXICAUX DANS LA LANGUE ROUMAINE AU XIX^{ÈME} SIÈCLE.

Une langue, quel qu'elle soit, ne suit pas toujours une évolution droite, un trajet linéaire, mais par suite des circonstances socio-historiques, politiques et culturelles et notamment grâce à ses utilisateurs, elle peut avoir des moments de régression, de stagnation ou d'avancement. Ces phases, ces étapes présentes dans l'évolution historique d'une langue peuvent être aussi bien décelés par l'analyse comparative des divers ouvrages lexicographiques parus au cours du temps, car, grâce au matériau lexical enregistré, « la mémoire sémantique » d'une communauté, d'un peuple, d'une époque est bien conservée et transmise le long des siècles. Dans ce cadre, les dictionnaires acquièrent une valeur symbolique particulière devenant ainsi des outils linguistiques et culturelles qui servent de base à la pénétration dans une langue donnée de nombreux mots « étrangers ». Dès le XVIII^{ème} siècle et tout au long du XIX^{ème} siècle, le roumain a emprunté du lexique au latin médiéval et aux langues néo-latins, aussi bien pour la langue de la vie quotidienne que dans les langues de spécialité. Les langues de spécialité et la terminologie deviennent ainsi des « sas d'entrée » des emprunts dans la langue commune. Ces termes issus d'autres systèmes linguistiques et culturels dénommant dans la plus part de situations des nouvelles réalités qui ont dû plus au moins s'adapter (d'un point de vue graphique, phonétique, morphologique et/ou sémantique) au système linguistique de la langue d'accueil ont accaparé l'attention et la réflexion des lexicographes. Dans notre exposé, nous nous proposons de dresser un panorama du traitement de l'emprunt dans les dictionnaires explicatifs roumains du XIX^{ème} siècle. L'examen du traitement de l'emprunt nous apportera des informations précieuses relatives à la présence ou à l'absence d'un marquage de l'emprunt, d'une part, et, d'autre part, nous permettra d'en saisir la dynamique de la langue roumaine au cours du XIX^{ème} siècle. Pour en mieux saisir cette dynamique, nous étudierons les vocables liés au domaine de l'astronomie. Dans ce sens, notre corpus choisi vise deux ouvrages lexicographiques tout à fait singuliers pour l'espace roumain ; il s'agit, d'une part, du Lexicon de Buda (LBe), un dictionnaire rédigé en quatre langues (roumain, latin, allemand et hongrois), paru à Buda en 1825 grâce aux efforts intellectuels conjugués de plusieurs érudits de l'École latiniste de Transylvanie et, d'autre part, du Vocabularu romano-francesu [Vocabulaire roumainfrançais] élaboré par Ion Costinescu et paru en 1870 à Bucarest. En nous appuyant sur ce corpus lexicographique, nous envisageons, dans un premier temps, de dresser un inventaire des vocables de l'astronomie, inventaire établi en vertu des définitions lexicographiques, et ensuite d'en comparer les résultats obtenus au but d'en saisir si, dans la dynamique de la langue roumaine du XIX^{ème} siècle, le lexique de l'astronomie subit ou non des changements (lexicaux et sémantiques). À la fin, nous allons avoir une perspective sur la dynamique de la langue roumaine, sur la circulation de ces vocables de l'astronomie au cours de presque cinquante ans. Comme dans d'autres langues, dans le cas du roumain aussi le lexique de l'astronomie subit en général peu de changements ; les vocables liés au domaine de l'astronomie pénètrent avec une certaine difficulté et il est rare que des mots déjà existants soient remplacés par de nouveaux mots appartenant au même champ sémantique. Ainsi, on a pu constater qu'il y a des mots stables, fixés dans la langue, et des mots instables qui entrent dans la langue, y restent pendant une certaine période et ensuite disparaissent totalement (c'est le cas surtout des mots employés au niveau régional), étant remplacés ou non par d'autres mots nouveaux (c'est le cas des mots qui se trouvent en concurrence synonymique, un mot ancien et un emprunt au latin médiéval ou à une langue romane). Cette constatation permet d'ouvrir de nouvelles perspectives d'analyse de la dynamique de la langue roumaine et par extension d'autres langues du monde. Pour en conclure, notre exposé dresse un panorama de l'évolution de la notion d' « emprunt » dans la langue roumaine au cours du XIX^{ème} siècle, d'une part, et, d'autre part, sur la circulation des vocables appartenant au domaine de

l'astronomie, tout en mettant en évidence leur langue d'origine et leur degré d'intégration totale ou partielle au système linguistique du roumain. Par notre approche, nous tenterons d'inscrire la problématique de l'emprunt linguistique dans le cadre plus large de l'histoire des idées linguistiques européennes.

ALFIERI, Gabriella & SARDO, Rosaria & BIFFI, Marco : VIVER: VOCABOLARIO RETICOLARE DELL'ITALIANO VERISTICO.

Tra le banche dati attinenti alla testualità letteraria archiviate in biblioteche digitali disponibili in rete con libero accesso (BIBLIOTECA ITALIANA, CEOD Corpus Epistolare Ottocentesco Digitale ecc.) o acquisibili a pagamento (BIZ - Biblioteca Italiana Zanichelli), manca un corpus unitario che documenti organicamente la produzione testuale del realismo italiano e ne consenta la comparazione con corpora omologhi in ambito romanzo ed europeo. Solo l'inserimento in un quadro sovranazionale e la prospettiva di una lettura trasversale potrà garantire un'interpretazione globale della narrativa realista europea, declinata con inevitabili specificità dovute ai singoli contesti culturali (naturalismo in Francia, verismo in Italia, socio-realismo in Inghilterra) e una comprensione dell'onda lunga dei naturalismi anche nell'ambito della paraletteratura. Il primo passo per una così rilevante e ormai irrinunciabile impresa euristica – che Centre Zola e Fondazione Verga stanno perseguendo in sinergia – non può che essere la costituzione di una banca dati attendibile e caratterizzante, che funga da macrotesto di riferimento per storici della letteratura e storici della lingua. Uno spoglio linguistico esauriente e condotto secondo le più attuali tendenze della lessicografia informatizzata costituirà uno strumento essenziale per accertare il testo delle opere, leggere e ordinare le correzioni dei manoscritti, qualificare e datare l'ingresso nella lingua nazionale di un patrimonio lessicale e fraseologico ancora in gran parte ignorato o sottovalutato. L'unico precedente in questa direzione è costituito dalle concordanze de I Malavoglia elaborate nei primi anni Ottanta dall'Accademia della Crusca e generosamente donate in forma di tabulati alla Fondazione Verga. Un analogo tentativo sperimentale fu svolto nello stesso periodo in collaborazione tra la Fondazione Verga e la Scuola Normale Superiore di Pisa per scrutinare sul piano della microsintassi il romanzo *Eva*. Questi tentativi pionieristici hanno trovato attuazione e continuità nel progetto qui presentato. Il VIVER (Vocabolario Reticolare dell'Italiano Veristico), ideato dalla Fondazione Verga nel 2015 e proficuamente avviato nel 2017 grazie a un finanziamento del MIBACT, e a una convenzione con l'Accademia della Crusca, mira a colmare una vistosa lacuna nella storia dell'italiano letterario e dell'italiano contemporaneo in generale. L'iniziativa si ricollega idealmente all'autorevole vocazione lessicografica della linguistica computazionale italiana (Biffi 2014, Biffi 2016 e Biffi, in stampa), e realizza la vocazione prioritariamente semantica della lessicografia (Avalle 1979). L'apporto di Verga e dei veristi al rinnovamento e all'arricchimento dell'italiano postmanzoniano è stato sancito da autorevoli studi teorici (Nencioni 1998), che attendono tuttora un supporto descrittivo. Interventi recenti hanno peraltro rimarcato la lungimiranza di progetti di politica culturale dello Stato postunitario, come quello di Pacifico Valussi (1864) o di Vincenzo Pasquini (1863) che anticipavano la ben più incisiva proposta manzoniana. Al drastico monocentrismo fiorentino poi adottato dall'autorità governativa, si contrapponeva un'unificazione linguistica 'morbida', fondata sulla progressiva convergenza 'federativa' ma toscano-centrica di esperienze letterarie, filologiche e lessicografiche di origine regionale (Alfieri 2011). Verificare i risultati di tale progettualità nella scrittura verista, in prospettiva nazionale ma sviscerandone le singole realizzazioni diatopiche, può costituire di per sé un fattore di

originalità e opportunità indagativa. La ricerca, istituzionalizzata dalla Fondazione Verga, abbraccia un orizzonte geoletterario e geolinguistico variegato che da Piemonte, Lombardia, Veneto e Friuli, attraverso Toscana, Abruzzo, Campania, Lucania e Calabria, si estende all'ambito insulare di Sardegna, Corsica, e Malta, per fermarci al contesto italo-romanzo. La Sicilia in questo caso costituisce ovviamente il centro di irradiazione della produzione testuale, mentre il modello linguistico rimaneva un toscano aperto a interferenze idiomatiche panregionali. L'indagine ricognitiva della testualità, di cui si danno qui di seguito le coordinate, ha avuto come obiettivo primario la predisposizione di un CORPUS DEI CORPORA DELLA LETTERATURA PRE- E POST-VERISTA, su cui si sta basando il vocabolario "reticolare" dell'italiano veristico, seguendo le linee teorico-metodologiche già tracciate per il Vocabolario dinamico dell'italiano contemporaneo, di cui si dirà più avanti. Al fine di consentire un elevato livello di fruibilità, si sta procedendo alla messa in rete di versioni digitalizzate di testi rappresentativi del patrimonio testuale dei veristi italiani, il cui "canone" è stato costruito sulla base del Congresso su "I verismi regionali" promosso nel 1992 dalla Fondazione Verga (AA.VV. 1996). Il corpus inclusivo della testualità verista italiana, di cui un campione rappresentativo è già consultabile in rete sul sito della Fondazione Verga (www.fondazioneverga.it), mira a incrementare la conoscenza della testualità postunitaria, e verista in particolare, nel suo complesso (autografi, edizioni originali e produzione critica o pubblicistica) in due direzioni: a) comunità dei lettori con finalità di divulgazione culturale; b) comunità scientifica con finalità di avanzamento delle conoscenze in ambito filologico-critico, storiocritico e storico-linguistico. In prospettiva europea si assumerà la più rappresentativa esperienza del verismo italiano come caso esemplare di studio per confrontarne gli intenti di rappresentazione estetica e le risultanze di messa in forma linguistica con gli intenti e le risultanze di scrittori naturalisti, in primis Zola e, per i contesti tedesco e inglese, Auerbach e, rispettivamente, Hardy. Nel corpus saranno incluse, a tal fine, le traduzioni di Verga in francese e di Zola in italiano, per rintracciare modalità 'accomunanti' di resa espressiva. Il progetto si fonda su solide basi filologiche, assumendo come testi di riferimento le edizioni originali, in massima parte già presenti nella biblioteca della Fondazione Verga, e si articola in ben definiti generi testuali, accorpati per ambiti geolinguistici: 1) novelle e romanzi di Verga, De Roberto e Capuana; 2) novelle e romanzi di autori realisti (es. Salvatore Farina, Roberto Sacchetti, ecc.) o veristi (De Marchi, Fogazzaro, Mezzanotte, ecc.); 3) teatro "borghese" e teatro verista dell'Otto-Novecento; 4) novelle, romanzi e teatro di autrici rappresentative del realismo e del verismo (Caterina Percoto, Neera, Serao, Deledda, Giselda Fojanesi, Evelina Cattermole, Contessa Lara, la Marchesa Colombi, o in ambito novecentesco Maria Messina, Adelaide Bernardini, Amalia Guglielminetti, Ada Negri, Annie Vivanti), isolatamente valorizzate in repertori antologici dedicati (Davico Bonino 1992) 5) letteratura per l'infanzia di tematica naturalistica (Capuana, Anna Vertua Gentile, Sofia Bisi Albini, De Amicis, Emma Perodi, Vamba) Come corpus di riscontro si assumeranno le traduzioni italiane di Zola e di Auerbach. Il patrimonio testuale dei veristi e di autori e autrici "minori" di Otto e Novecento, articolato nei vari generi e sottogeneri sopra indicati, può costituire un importante corpus da integrare ad altri corpora già costituiti, per arrivare a un'interrogazione unificata, con adeguato metamotore, sulla falsariga seguita dall'OVI per il TLIO (Tesoro della Lingua Italiana delle Origini) e dall'Accademia per le varie "Crusche" e per le fonti testuali che costituiscono la banca linguistica della Crusca (Biffi 2016, Biffi 2018, Biffi in stampa). Questa concezione lessicografica porta alla trasformazione del vocabolario in sistema lessicale digitale come il DWDS (Digitales Wörterbuch der deutschen Sprache) della Brandenburgische Akademie der Wissenschaften. Va sottolineato innanzitutto che, come già accennato, il VIVER confluirà in parte nel VODIM (Vocabolario dinamico dell'italiano Moderno), oggetto di due PRIN

2012 e 2015 (Progetto di Interesse Nazionale), di cui il P.I. è il Presidente dell'Accademia della Crusca, prof. Claudio Marazzini, e di cui Gabriella Alfieri, Presidente del Consiglio Scientifico della Fondazione Verga, coordina l'unità di ricerca dell'Università di Catania (Alfieri – Alfonzetti – Sardo 2016). Le fasi e le potenzialità del progetto sono così sintetizzabili: - progettazione e realizzazione di un'applicazione web ad accesso pubblico con tecnologia Angular2 e Typescript sul lato client, e una API .NET sul lato server; - realizzazione di un database di "catalogo" delle opere contenute; - consolidamento di un indice dei testi, per consentire ricerche essenziali, sul tipo dell'indice implementato per www.tommaseobellini.it elaborato per l'Accademia della Crusca; - definizione del dialetto XML a base TEI e progettazione software per convertire i file word nel dialetto XML a base TEI - attribuzione di opportuni metadati ai testi, per la creazione di un sistema di indicizzazione e ricerca (in particolare realizzazione di un'applicazione web su lato client e un "backend" API su lato server che rendano consultabile il corpus sia per via di ricerca sistematica che per via di libera esplorazione (lettura e simili). - impostazione di varie sezioni tematiche di approfondimento, tipiche dei testi "statici" (romanzi, novelle, opere teatrali con relativi paratesti, quali introduzioni, didascalie, ecc.). Sarà così possibile predisporre il corpus per un sistema di interrogazione interna (finalizzata alla costruzione di altre risorse) ed esterna (portale accessibile al largo pubblico), per realizzare l'obiettivo di una pubblicazione on line del lessico dei veristi. I risultati attesi presentano un'indiscussa valenza sociale e culturale, con enormi potenzialità anche sul piano didattico.

ÁLVAREZ VIVES, Vicente : La impronta de la obra herediana en el estudio del léxico hispánico a partir del *Diccionario del Español Medieval (DEM)*.

Es indiscutible la importancia del siglo XIV como etapa de «plenitud» del léxico medieval (tanto patrimonial, como culto). Con los antecedentes textuales de la Escuela de Traductores de Toledo, la gran labor de Alfonso X (siglo XIII), y los referentes literarios contemporáneos como el Libro de Buen Amor o El Conde Lucanor, por citar algunos de los más conocidos, queda patente que el siglo XIV pueda ser entendido como una etapa de consolidación del léxico –y, por extensión, de la lengua– medieval. El gran maestro D. Juan Fernández de Heredia (ca. 1310-1396) fue un aragonés muy influyente durante el siglo XIV que “desplegaría una gran labor cultural, rodeado de copistas (Bernardo de Jaca), traductores (Domingo García Martín, fray Nicolás) e iluministas que produjeron libros de gran calidad y gusto delicado, en una línea que le hace comparable a Alfonso el Sabio” (Cortés Arrese 1983: 327). Contamos con una considerable bibliografía sobre Fernández de Heredia con estudios dedicados, por ejemplo a la persona (Montaner Frutos 1996), a su función como traductor (Adelino Rodríguez 1985 y 1986, Romero Cambrón 2011), a su lengua y estilo (Badia i Margarit 1996-1997). Concretamente, se pretende observar el léxico empleado en sus obras e indicar la influencia que tuvo en su época y que queda demostrada en todas las obras del corpus herediano que han sido consultadas en la elaboración del Diccionario del Español Medieval (DEM). El quehacer lexicográfico conlleva tareas interesantes como la de describir nuevos vocablos o antiguos valores de voces usadas en la actualidad, pero también comporta tareas mecánicas como las del cómputo de fichas, revisión y comprobación de datos, rastreo de posibles errores o de hipotéticas confusiones, entre otros. A este propósito, durante el tiempo de recuento de fichas y de preparación de materiales del fichero del DEM para el proyecto DEMel (Diccionario del Español Medieval Electrónico), cayeron en mis manos varias gavetas repletas de fichas dedicadas a palabras como batalla o, un mejor ejemplo

es el de guerra, que, con un total de 4 gavetas –números 223 a 226–, por ejemplo la número 224 está íntegramente compuesta por fichas que recogen testimonios de las obras de Heredia. Sin lugar a dudas, cuando se pueda acceder a los materiales del fichero del DEM, gracias al proyecto DEMel, se podrá comprobar la totalidad de las referencias heredianas en el léxico del español medieval. Esto permitirá ampliar y completar la información, que, por el momento, se ve reducida al análisis de los datos recogidos en el DEM, que no son pocos. Concretamente, para este estudio han sido analizados los 3 últimos fascículos publicados (fasc. 24, 25 y 26) y han sido recogidas todas las entradas que contienen ejemplos de las obras de Fernández de Heredia; por citar algunas referencias, Libro de actoridades (Rams de flors), Cronica de los conqueridores I/II, Cronica de Morea, Cronica Troyana, Eutropius, Libro de Marco Polo, por ejemplo. En esta contribución se cotejará con otros diccionarios del español la existencia de las entradas y, en algunos casos particulares, se añadirán comentarios sobre eventuales cambios semánticos. El análisis se presentará tomando como base cada una de las obras de Fernández de Heredia que conforman parte del corpus del DEM. Se intentará establecer algunas conclusiones acerca de la naturaleza etimológica de las palabras, como por ejemplo helenismos, arabismos, aragonesismos.

AUTELLI, Erica & LUSITO, Stefano & KONECNY, Christine : Verso un nuovo repertorio di collocazioni ed espressioni idiomatiche genovesi. Problemi di classificazione, pronuncia e traduzione in *GEPHRAS*.

Per quanto stiano acquisendo sempre più importanza anche in Italia (cfr. Mollica 2015, 13), gli studi sulla fraseologia – e specialmente sulla fraseografia dialettale (cfr. rispettivamente Piirainen 2007, 538 e Autelli in svolgimento) – sono ancora ai primordi, tanto che per il futuro si auspicano ulteriori indagini circa l’attestazione dei fraseologismi nelle diverse varietà diatopiche della penisola (cfr. Cini 2005, 275). Sulla base di questi presupposti, ad agosto 2018 ha avuto inizio presso il Dipartimento di Romanistica dell’Università di Innsbruck il progetto «GEPHRAS: Das ABC genuesischer und italienischer Phraseme» (Autelli *et al.* in svolgimento), primo repertorio contrastivo genovese-italiano dedicato alla registrazione delle collocazioni e delle espressioni idiomatiche relative a un idioma locale d’area italiana. Più nello specifico, il lavoro si concentra sul «genovese» inteso non solo come varietà ligure più diffusa a livello regionale (sulla costa tra Bergeggi e Moneglia e in gran parte del relativo entroterra), ma anche come quella dotata di maggior prestigio, in virtù dell’impiego in ambito letterario e del ruolo storico detenuto dal capoluogo (cfr. Petracco Sicardi 1974; Toso 2002). L’unicità di un lavoro del genere ha comportato l’esigenza di compiere scelte ben definite a livello di metodologia. Il presente contributo verterà in particolare su tre aspetti che rappresentano vere sfide per la compilazione di tale repertorio: la struttura delle entrate, la resa fonica dei fraseologismi selezionati e le rispettive traduzioni. Il primo aspetto parte da un’ampia concezione di fraseologismo (cfr. ad es. Burger 2003; Valenčič Arh 2014): se da una parte è parso preferibile escludere l’ampio insieme di proverbi, in buona parte già documentati nel 1869 da Staglieno, si è ritenuto invece significativo includere, oltre alle espressioni idiomatiche, altre categorie fraseologiche quali le collocazioni e i fraseologismi strutturali e comunicativi (cfr. Burger 2003, 36). L’obiettivo è quello di attestare il maggior numero di combinazioni possibili, tanto a fini fraseografici quanto per favorire la preservazione del genovese, classificato dall’UNESCO (2010) a rischio di estinzione insieme alle restanti varietà regionali. Si tratta di una proposta lessicografica al momento unica del suo genere, data la generale scarsità di dizionari collocazionali bilingui, l’asciutaggine e l’obsolescenza di gran

parte dei repertori genovesi e la mancanza di studi fraseografico-dialettologici in merito ai singoli fraseologismi (cfr. Autelli in svolgimento). Oltre a riassumere la complessa suddivisione in categorie morfosintattiche delle entrate relative ai cento lemmi selezionati (per il momento, sostantivi iniziati per A, B e C, varianti diacritiche incluse), verranno segnalati aspetti rilevanti nell'ambito della linguistica delle varietà. Poiché il repertorio prevede l'inserimento di ampio materiale audio (liberamente consultabile sulla piattaforma *online* del progetto, insieme ai contenuti testuali), è parso poi necessario individuare un preciso modello di pronuncia che potesse essere giudicato «neutrale» da parte degli utenti, facendo attenzione a escludere persino i localismi meno riconoscibili per il parlante medio. La scelta è ricaduta su quella che un tempo era considerata la pronuncia «urbana» del capoluogo, sia perché quella più aderente al modello grafico consacrato nell'uso, sia perché l'unica a essere giudicata priva, da parte degli stessi parlanti, di taluni tratti che invece caratterizzano in senso ultralocale varietà anche prossime a quella selezionata. Una seconda parte dell'intervento riassumerà quindi non solo le caratteristiche fonetiche di questa specifica sottovarietà (la più fortemente interessata da processi di dinamismo all'interno del territorio regionale), ma anche le principali discrepanze che la separano da quelle a essa prossime: si tratta di differenze che arrivano talora a riguardare la stessa struttura sillabica delle parole, a causa della complessa evoluzione linguistica attraversata dal genovese a partire dalle prime epoche di attestazione (cfr. Parodi 1898, 1901 e 1902-1904; Ageno 1957; Toso 2004). Il terzo punto dell'intervento riguarderà le traduzioni italiane dei fraseologismi genovesi, ora tratte dai repertori locali già esistenti, ora fornite da dialettofoni attivi coinvolti nel progetto. In aggiunta al fatto che talune strategie di traduzione non possono essere applicate agli equivalenti di singoli fraseologismi, si metterà in evidenza come la qualità delle traduzioni riscontrate nei dizionari genovesi sia in genere molto bassa: si tratta di un dato che interessa anche i repertori più estesi e rinomati (come il Casaccia² 1876 e il Frisoni 1910), che oltre a trascurare una quantità assai rilevante di materiale ancor oggi diffuso nell'uso si rivelano piuttosto generici in merito all'attestazione di fraseologismi (non distinguendo, ad esempio, le collocazioni dalle espressioni idiomatiche). Si tenterà di classificare le problematiche riscontrate offrendo esempi concreti sia in genovese che in italiano. Al contrario di ciò che fanno molti traduttori, che spesso tendono a parafrasare anche qualora esista un equivalente (cfr. Zybatow 1998, 155), si cercherà di fornire i traduttori esatti salvo casi di 'equivalenza zero' (cfr. Korhonen 2007, 577 sg.). Nell'intervento verranno infine messe in discussione soluzioni personali di parlanti nativi di genovese che non sono probabilmente (ancora) condivise dalla comunità: si tratta di una riflessione che offre ulteriori spunti di ricerca, da affrontare (non esistendo ancora grandi *corpora* del genovese contenenti materiale fraseologico) tramite indagini condotte fra genovesofoni madrelingua. Con questo intervento si spera quindi di fornire un quadro complessivo di quelle che possono essere le maggiori difficoltà nella compilazione del repertorio in questione e di offrire un contributo sia alla ricerca fraseologica, sia a quella lessicografica, dialettale e traduttologica (nel territorio italiano e non solo). L'obiettivo è quello di mettere a disposizione un modello che possa essere adottato anche per altre varietà e che, oltre a generare nuovi impulsi per la fraseodidattica, sia anche in grado di proporsi come strumento di salvaguardia per lingue o dialetti a rischio di scomparsa.

BOHBOT, Hervé & FRONTINI, Francesca & FAUCHER, Alexandre : Le projet Nénufar, nouvelle édition numérique du *Petit Larousse illustré* (1906-1948).

Le Petit Larousse illustré est l'héritier d'une longue tradition éditoriale de la maison Larousse en matière de dictionnaires de langue française. Son ancêtre, le Nouveau Dictionnaire de la langue française, publié par Pierre Larousse (1817-1875) en 1856, se vend à plus de quatre millions d'exemplaires en un demi-siècle et introduit la célèbre présentation en trois parties : langue française, citations ("pages roses") et partie historique (Pruvost, 2006a). La première édition du Petit Larousse est publiée en 1905 sous la direction de Claude Augé, déjà directeur du Dictionnaire complet illustré (1889) et du Nouveau Larousse Illustré (7 volumes publiés de 1897 à 1904). Le succès éditorial est immédiat et plusieurs centaines de milliers d'exemplaires sont vendus chaque année (appelée "millésime"). Environ 230 tirages ont lieu entre 1905 et 1924 quand l'ouvrage est entièrement remanié et renommé Nouveau Petit Larousse illustré. Suivront environ 430 éditions entre 1924 et 1947, année qui verra une nouvelle refonte et le retour au nom Petit Larousse illustré. Les dictionnaires sont les « outils d'une langue et d'une culture » (Pruvost, 2006b) et le Petit Larousse, diffusé à plusieurs centaines de milliers d'exemplaires chaque année depuis plus d'un siècle (Dubot, 2012), présent dans la très grande majorité des foyers français a constitué un levier de la démocratisation du savoir (Cormier, Francoeur, 2005). Sa disponibilité sous forme numérique est donc très intéressante pour l'étude du français. Le nom du projet Nénufar (Nouvelle édition numérique de fac-similés de référence), est inspiré par la surprenante polémique récente sur une « réforme de l'orthographe » qui avait en fait presque 30 ans, amplifiée par les médias et les réseaux sociaux. Un des exemples les plus souvent cités par ses détracteurs, était la recommandation du remplacement de nénuphar par nénufar, qui était pourtant une graphie courante durant toute la première moitié du XXe siècle, comme l'atteste le Petit Larousse depuis 1905 (entrée « NÉNUFAR ou NÉNUPHAR »). Dans l'édition 1948, l'entrée devient « NÉNUPHAR ou NÉNUFAR » et dans l'édition 1955, la graphie nénufar disparaît et est oubliée au point de devenir choquante à certains aujourd'hui. On pourrait également citer à priori (avec accent), fiord, ognon, présents dans la première édition du Petit Larousse. Au-delà des variations orthographiques, la connaissance des premières éditions du Petit Larousse permet de remettre en évidence des évolutions relativement récentes de la phonétique ([distrik], [lokouass] pour district et loquace en 1906), des mots « nouveaux » (antimilitarisme en 1911, boche en 1917, etc.) et des changements de l'univers de référence, révélés par les articles, parfois amusants (à l'article AVIATION, en 1905 « on a fait de nombreuses tentatives [...] mais le problème n'est pas encore résolu », qui devient en 1911 « les avions ont victorieusement résolu le problème du plus lourd que l'air ») ou choquants (on peut lire jusqu'en 1951 à l'article NÈGRE « Leur crâne est dolichocéphale, leur face longue, leur nez écrasé, leurs lèvres sont grosses, etc. » et jusqu'en 1967 à l'article JUIVERIE « Rapacité sordide »). Ce dictionnaire populaire est ainsi une mine d'observations et d'enseignements, à travers ses prises de position et leur évolution, sur une société dont il se veut le reflet. En fait, la neutralité revendiquée par le Petit Larousse est toute relative, notamment à travers ces « exemples qui surgissent au détour de tel ou tel mot [qui] font en réalité partie d'une éducation indirecte de la population » (Pruvost, 2004). Plus généralement, « traquer les changements de définitions, minimes ou majeurs, ne manque pas d'intérêt, car il faut le rappeler, les dictionnaires millésimés reflètent d'assez près la doxa d'un moment sur tous les sujets et, à ce titre, offrent une sorte de référence précieuse. On y repère ainsi parfois avec un grand étonnement l'évolution de telle ou telle représentation à partir d'un mot qui nous semble sans surprise » (Pruvost, 2015). Il y aurait matière à de nombreuses études sociolinguistiques, ce qui représentait une entreprise à bien des égards

difficile : malgré ou à cause de sa popularité, les ressources numériques concernant le Petit Larousse et plus généralement lexicographie de la première moitié du XXe siècle sont extrêmement rares et peu exploitables. Les quelques éditions disponibles, scannées en mode image et en basse résolution (Google Books, Gallica), ne permettent que la lecture. Une édition numérique de la partie langue française de la première édition du Petit Larousse a toutefois été entreprise en 2004 et publiée en 2009 (Manuélian et al.). Malheureusement les données n'étaient pas en libre accès, le projet n'a pas été maintenu et son site web n'est plus opérationnel depuis 2015. Le projet Nénufar se veut une refonte et un approfondissement de ce premier projet de numérisation du Petit Larousse en y intégrant l'ensemble des éditions successives libres de droits (en l'occurrence, celles publiées il y a plus de 70 ans) afin de pouvoir étudier son évolution. Pour respecter l'unité et la structure de l'ouvrage, il était également indispensable d'ajouter la partie historique, liée par plusieurs centaines de renvois croisés à la partie langue. Nous visons ainsi la publication des données de chaque millésime de 1906 à 1948 avec une datation annuelle des changements. Les éditions les plus importantes ont été numérisées (première édition, éditions précédant et suivant les refontes). Après reconnaissance optique des caractères et correction, les données textuelles ont été balisées selon la norme TEI (Bański et al., 2017) qui est la référence retenue dans le cadre de plusieurs projets de rétronumérisation de dictionnaires historiques, notamment le projet Basnage (Williams & Galleron, 2016). L'exploitation des données du projet pour la recherche en sciences humaines et sociales et notamment pour la linguistique et les technologies du langage, peut se faire par trois voies d'accès : - une plateforme web, <http://nenufar.huma-num.fr>, proposant une interface de consultation simple et ergonomique sur tous supports numériques qui intègre l'ensemble des fonctionnalités attendues d'un dictionnaire électronique : renvois hypertextes généralisés, recherche en texte intégral d'un lemme et de ses formes dérivées dans l'ensemble du corpus, requêtes multicritères par marqueurs (étymologie, niveaux de langue, domaine...), requêtes diachroniques permettant d'identifier les changements d'articles (entrées apparues, modifiées ou disparues telle année ou durant telle période, mise en évidence des parties de textes modifiées, etc.), consultation des documents originaux en haute définition. - le téléchargement, en vue d'une utilisation experte, des sources XML déposées sur la plateforme Ortolang, qui garantit par ailleurs la préservation à long terme des données. - l'interrogation en ligne des données dans le format du web sémantique selon le modèle OntolexLemon (Cimiano et al. 2016, Bohbot et al. 2018). La publication des données a commencé en juillet 2018 et se poursuivra régulièrement au cours des années 2018 et 2019. La méthodologie et les outils développés pourront par la suite être appliqués à d'autres dictionnaires patrimoniaux.

CAMPOS SOUTO, Mar : La ordenación de las acepciones en el NDHE: bases para el establecimiento de los criterios lógico-lingüístico y temático.

En distintos diccionarios diacrónicos (y, en particular, en la lexicografía histórica del español), la ordenación de las acepciones de un artículo se ha ajustado a dos criterios: el cronológico y el lógico-lingüístico (véanse, a modo de ejemplo, el DHLE, el OED y el TLF), en tanto que repertorios de orientación sincrónica han optado, en ocasiones, por la frecuencia como pauta ordenadora de las diversas acepciones de un vocablo. Los denominados orden histórico (cronológico), orden empírico (frecuencia) y orden lógico han determinado, pues, en buena medida la prelación de las acepciones en los diccionarios publicados en papel. Como ya advertía Lapesa en el prólogo al primer tomo del DHLE de la Real Academia Española, es ciertamente habitual que “la sucesión reclamada por el orden de primeras apariciones” esté “en pugna con el encadenamiento originario que es preciso

suponer entre unos significados y otros” (1972: XI). Es bien sabido que en el DHLE se distinguen, precisamente, series, acepciones y subacepciones (“divisiones semánticas de distintos grados”) en función de un criterio semántico. Montes, con respecto al DCR de Cuervo, ya indicaba que el orden lógico, si bien presenta un carácter un tanto vago, “puede tomarse como equivalente del orden en que la investigación histórica demuestra que se produce de ordinario la evolución semántica: de lo concreto a lo abstracto, de lo material a lo inmaterial, etc.” (1991: 8; véase también Imbs, 1971: XXV). Si dirigimos nuestra mirada a la clasificación temática de las acepciones, comprobaremos que con frecuencia esta se añade en una fase avanzada de la construcción del repertorio e, incluso, como un proyecto independiente; sirva como ejemplo el *Historial Thesaurus* del OED (proyecto llevado a cabo en la Universidad de Glasgow, que se presenta como “a taxonomic organization of all the contents of the OED”). El surgimiento, en las últimas décadas, de la lexicografía electrónica (que ha supuesto una transformación radical de los métodos de trabajo) ofrece como indudable ventaja la posibilidad de ofrecer distintos modos de presentación de los datos, modos que, por su parte, requieren una rigurosa formalización previa y un desarrollo informático que facilite la publicación dinámica de las informaciones. Por consiguiente, dotar de la posibilidad de consultar los artículos de un diccionario en virtud de un criterio lógico-lingüístico o de un criterio temático supone una operación previa de análisis y de establecimiento preciso de las pautas que guiarán esa ordenación. El Nuevo diccionario histórico del español (NDHE) de la Real Academia Española, un diccionario nativo digital y en curso de elaboración, en la actualidad permite ordenar las acepciones de los artículos publicados según dos criterios: el cronológico y el de frecuencia (para este diccionario, véase Campos Souto 2015 y 2016 y Salas Quesada y Torres Morcillo, 2011 y 2015). Por defecto, las acepciones se organizan en virtud de un criterio cronológico (en primer lugar figura la más antigua), si bien los usuarios tienen la posibilidad de optar por presentarlas por su frecuencia absoluta (un criterio que, no obstante, presenta problemas, pues está estrechamente ligado a la selección de las fuentes documentales que suministran los testimonios en que se sustenta el análisis de las palabras). Uno de los retos que se presentan al NDHE en el futuro inmediato es el de poder ofrecer otros dos modos de acceso a las acepciones de cada artículo: el temático y el lógico-lingüístico. Esta comunicación persigue presentar, a partir del análisis de los artículos publicados del NDHE, los criterios que guiarán esos dos modos de organización de las acepciones y que servirán de pauta para el desarrollo informático; dado que el repertorio seguirá creciendo y evolucionando, estas pautas han de considerarse necesariamente provisionales (esto es: servirán para construir el primer prototipo de esa sección de las herramientas de redacción y visualización del diccionario). Para el primer criterio ordenador (el temático), se optará por el establecimiento gradual de taxonomías, que se irán construyendo a medida que se elaboran los diferentes artículos; en la comunicación se mostrarán los primeros resultados que facilita ese modo de organizar las acepciones. Por otra parte, para la aplicación del orden lógico-lingüístico, se presentarán las conclusiones obtenidas del análisis de la aplicación de criterios que pueden ser relevantes para este asunto, como los siguientes (entre otros posibles): — Criterio etimológico: en el caso de vocablos españoles procedentes de otras lenguas (latín, francés, inglés, etc.), distinción entre acepciones etimológicas y acepciones no etimológicas (y limitaciones para establecer, en ciertos casos, el carácter etimológico o no de las acepciones analizadas). — Criterio morfológico: en el caso de voces formadas por procedimientos de formación de palabras, se examinarán, entre otras, las siguientes cuestiones: a) prelación de las acepciones de la palabra base; b) prelación de los valores de los afijos; c) posibilidad de incorporar criterios categoriales. — Criterio semántico: Relaciones de inclusión y mecanismos de cambio semántico; conveniencia (o no) de establecer prelaciones entre distintos mecanismos. Interacción con el criterio

etimológico. Por último, se analizará la relación entre el criterio lógico-lingüístico y el criterio cronológico para determinar si realmente nos hallamos ante pautas organizadoras totalmente independientes.

CANO GONZÁLEZ, Ana María : La lengua asturiana en los diccionarios etimológicos panrománicos: Del REW al DÉRom pasando por el DELLA (*Dictionariu etimolóxicu de la llingua asturiana*).

En la presente comunicación pretendo hacer un análisis de la presencia y tratamiento de la lengua asturiana en los diccionarios etimológicos panrománicos, es decir, en el *Romanisches Etymologisches Wörterbuch* (REW) de Meyer Lübke y en el *Dictionnaire Étymologique Roman* (DÉRom). Primeramente se revisarán los aspectos fundamentales que separan ambos diccionarios y que no solamente tienen que ver con la lejanía en el tiempo de la elaboración de ambas obras, aunque sin duda pueden estar relacionadas. Algunas de estas diferencias son más bien de carácter formal: autoría individual (REW) / colectiva (DÉRom), lengua de redacción (alemán en el primer caso / una lengua romance –francés- en el segundo), número de lenguas manejadas (unas pocas en el REW frente a la casi totalidad de las mismas en el DÉRom), etc. Otras son más de fondo y tienen que ver con la nueva concepción de la etimología. Y aquí radica una de las principales innovaciones del DÉRom, la reconstrucción del étimo a partir de la comparación de las lenguas romances teniendo en cuenta las dos caras del signo lingüístico, el significante y el significado; poniendo especial cuidado en la fijación de las primeras documentaciones en las distintas lenguas romances de modo que se pueda reconducir el étimo establecido hasta los resultados romances actuales. No se puede olvidar la aportación del DÉRom en el conocimiento de las diferentes áreas románicas resultado en buena medida de la distinta cronología del desgajamiento de las diferentes partes de lo que fue el Imperio Romano. A continuación se revisará la presencia, ciertamente escasa, de la lengua asturiana en el REW. En tercer lugar se analizará la presencia del asturiano en el DÉRom donde es una de las lenguas de cita obligada, gracias justamente a que el asturiano cuenta con un diccionario etimológico, el *Dictionariu Etimolóxicu de la Llingua Asturiana* (DELLA) de García Arias, del que, aunque el primer volumen (A-B) se publicó en este mismo año de 2018, su autor puso siempre generosamente a disposición de los redactores del DÉRom los correspondientes artículos del DELLA, con independencia de su carácter inédito o no. Un avance de lo que finalmente es el DELLA puede verse en los 5 volúmenes de *Propuestes etimolóxicques* del autor publicadas por la Academia de la Llingua Asturiana. Finalmente, tras un breve resumen de los avances realizados en las últimas décadas en el ámbito de la filología asturiana, se analizará la que es por el momento su obra cumbre, el DELLA (lematización, estructura de los diferentes artículos, soluciones modernas y distribución diatópica, documentación histórica con especial atención a la primera documentación conocida, estudio etimológico realizado, etc.) y su aportación a la filología hispánica (comparación con los diccionarios etimológicos de las lenguas de dicho ámbito románico) y románica (especialmente en referencia al DÉRom y al proyecto PatRom).. El DELLA, del que en estos momentos (septiembre de 2018), como ya se dijo, solo hay publicado un volumen, constará de siete tomos, de los cuales el II (C-D) y el III (E-F) están a punto de entrar en la imprenta con lo que previsiblemente verán la luz en 2019; y los otros (IV, letras G a O; V, letras P a R; VI, letras S a Z y bibliografía; y VII, dedicado a los índices), sin poder precisar la fecha, se publicarán en un futuro próximo.

CARRASCO, Inés & CARRASCO, Pilar : Aspectos léxicos en documentación colonial americana relacionada con los Procesos de fe.

El trabajo que proponemos se inscribe dentro de las acciones realizadas en el proyecto Red para un Corpus Diacrónico del Español en el Mundo (RECORDEM, IP: Pedro Sánchez Prieto), cuyo objetivo primordial es abordar la problemática americana del documento archivístico e incorporar un fondo básico de documentos que permita el estudio histórico a partir de materiales de primera mano. Los documentos que analizaremos en este estudio han sido transcritos de acuerdo con los criterios filológicos recomendados por la RED Internacional CHARTA y están clasificados en el apartado de “grupos sociales”. Recogen testimonios interesantes de carácter coloquial por lo que se solventa el principal problema con el que se han enfrentado los estudios del español americano. Están integrados, entre otros, por expedientes relacionados con la Inquisición en América de los siglos XVII y XVIII, sobre los casos de mujeres acusadas de brujería y pertenecientes a los fondos del Archivo General de Indias y Archivo Histórico Nacional. Nos proponemos ofrecer una breve muestra de las posibilidades que para el estudio del léxico permite el corpus a través del comentario de una serie de términos propios del español de Perú y México. Las diferentes partes que forman el documento nos permitirán establecer distintas tradiciones discursivas caracterizadas con un léxico más popular, propio de la oralidad, en la parte de la llamada probanza o en las declaraciones de testigos (apodos, oficios, nombres populares de las partes del cuerpo, enfermedades, hierbas medicinales, formaciones con sufijos derivativos con significado diminutivo, onomatopeyas, etc), frente al léxico más formal y formulístico que se observa en el desarrollo de la sentencia realizada por los fiscales o inquisidores. Esta propuesta se puede encuadrar dentro de la sección Lexicografía, en el subapartado de Lexicografía de corpus. El análisis de los corpus documentales es una de las principales vías para el hallazgo de voces poco usadas o restringidas a espacios geolectales específicos. Igualmente, está permitiendo conocer diferentes acepciones y usos de voces no recogidas en las fuentes lexicográficas y que podrán ser tenidas en cuenta en el diccionario histórico del español.

CLAVERÍA, Gloria & JULIÀ, Carolina : Los italianismos en el español y en el catalán contemporáneos.

Esta propuesta de comunicación pretende desarrollar un estudio de los neologismos procedentes del italiano más recientemente incorporados tanto en el español como en el catalán (finales de siglo XX y principios de siglo XXI). El estudio se integra en el proyecto internacional de investigación Observatorio de los italianismos en el mundo (OIM) de la Accademia della Crusca (Heinz 2017), dirigido por Luca Serianni y Matthias Heinz, en el que participan investigadores de un buen número de universidades (Uniwersytet Warszawski, Universidade de Lisboa, Universidad de Sevilla, Universität Salzburg, Universitat Autònoma de Barcelona, Università della Calabria, Jagiellonian University, Eötvös Loránd University). Este proyecto tiene como punto de partida el material del Dizionario di italianismi in inglese, francese e tedesco (DIFIT 2008) dirigido por Harro Stammerjohann, un diccionario que será consultable en línea (<http://www.italianismi.org>) y que actualmente se encuentra en proceso de revisión y ampliación con la recolección de los italianismos presentes en otras lenguas europeas, tanto románicas (español, catalán y portugués) como no románicas (húngaro y polaco). Se desarrolla mediante una plataforma informática que homogeneiza la obtención y gestión de los datos a la vez que permite y facilita la comparación interlingüística, una metodología que proporciona grandes beneficios a la investigación lexicológica. En el caso del

español, se han tomado como punto de partida inicial para el acopio de los italianismos diversas fuentes lexicográficas entre las que se encuentran el Diccionario crítico etimológico castellano e hispánico (DECH) y los diccionarios consultables en el NTLLE (2001); y, hasta el momento, la compilación de datos se constriñe al español peninsular. En el caso del catalán, la recopilación de italianismos se fundamenta en el diccionario de Gómez Gane (2012) y el Banco de neologismos BOBNEO, consultable en línea en el Centro Virtual Cervantes (https://cvc.cervantes.es/lengua/banco_neologismos/) procedente de las bases de datos de neologismos del Observatori de Neologia del IULA (<http://www.iula.upf.edu/>). Tomando como base los trabajos que se llevan a cabo dentro del proyecto OIM, la comunicación tendrá como objetivo el análisis de los italianismos de incorporación reciente (finales de siglo XX y principios de siglo XXI) y sus principales particularidades (cfr. Rainer 2005, Feliu 2008, Veny 2008, Álvarez de Miranda 2009, Dworkin 2012, Pistolesi 2013, Cabré et al. 2014, Gómez Gane 2016, Massanell 2016, Kamp 2016). Se intentará determinar las características evolutivas que presenta el préstamo del italiano en el español y catalán contemporáneos desde el análisis de varios aspectos cruciales para su comprensión: a) la variación formal que presenta y su integración en los diccionarios; b) los campos semánticos a los que pertenece y los valores semánticos adoptados; c) el comportamiento morfológico del italianismo (flexión y derivación); d) la comparación con el comportamiento de otras lenguas románicas y del propio italiano como lengua donante; e) los procesos de préstamos directo o préstamo indirecto en el que interviene otra lengua en el proceso de transmisión al español; f) contraste con el uso en los corpus.

COLUCCIA, Rosario & MANNI, Paolo : *Impostazione, metodo e finalità di una nuova impresa lessicografica : il Vocabolario Dantesco*

Progetto. A partire dal settembre 2015 l'Accademia della Crusca e l'OVI (Opera del Vocabolario Italiano – Istituto CNR) hanno avviato i lavori di redazione del «Vocabolario Dantesco» (VD), impresa lessicografica di vasto respiro che si inserisce nelle iniziative scientifiche attivate in tutto il mondo in occasione del settimo centenario della scomparsa di Dante (2021). Il VD, destinato a raccogliere l'intero patrimonio lessicale contenuto nelle opere di Dante (volgari, a partire dalla *Divina Commedia*, e latine), sarà consultabile liberamente in rete, quindi sempre aggiornabile e ampliabile. È inoltre prevista, con gli opportuni adattamenti, una pubblicazione cartacea finale.

Organizzazione. La Commissione scientifica è costituita da Paola Manni, Giancarlo Breschi, Rosario Coluccia, Giovanna Frosini, Lino Leonardi, Aldo Menichetti e Mirko Tavoni.. A questo nucleo si affiancano, anche allo scopo di favorire il collegamento tra la nuova impresa lessicografica e il TLIO, Rossella Mosti e Zeno Verlati; è stato inoltre chiamato a dare il suo contributo al progetto Giuseppe Marrani. Il lavoro di prima redazione dei lemmi è svolto da Francesca De Blasi, Barbara Fanini, Cristiano Lorenzi Biondi, Fiammetta Papi, Veronica Ricotta; prima di pervenire alla pubblicazione le voci sono sottoposti al filtro collettivo e alla approvazione di tutti i componenti del VD. Il sito è curato da Salvatore Arcidiacono.

Metodo. La priorità accordata alla versione informatica si lega a precise ragioni metodologiche. La realizzazione del VD trova il suo naturale fondamento negli stessi strumenti che sono alla base del lavoro lessicografico dell'OVI e si sviluppa in parallelo al *Tesoro della Lingua Italiana delle Origini* (TLIO). Nel VD l'impostazione lessicografica tradizionale – che si prefigge di precisare il valore semantico dei lessemi danteschi attraverso definizioni puntuali e sintetiche – viene affiancata e significativamente arricchita attraverso un

dialogo continuo con altre importanti risorse informatiche, fondamentali sotto il profilo storico-linguistico o sotto quello esegetico-interpretativo. All'interno di ogni singola Scheda lessicografica collegamenti ipertestuali diretti consentono l'accesso agli articoli corrispondenti del TLIO, della [Lessicografia della Crusca in rete](#), della versione informatizzata dell'[Enciclopedia Dantesca](#) (ED) e al corpus dei Commenti danteschi. **Fasi.** Il VD prende avvio dalla *Commedia*, opera centrale nell'ambito della produzione dantesca e dell'intera storia culturale italiana ed europea. La scelta di avviare il cantiere del VD proprio dal poema e la necessità di doversi misurare, da subito, con la delicata questione della scelta dell'edizione di riferimento, ha rafforzato l'idea di dar vita a uno strumento che rispondesse a una concezione più aperta del lessico che ruota attorno all'opera dantesca e alla sua ricezione. Le voci del VD segnalano sistematicamente (quando esistono) le varianti significative testimoniate dalla tradizione manoscritta, desunte dall'apparato dell'edizione Petrocchi e da ogni altra fonte utile. Quell'apparato offre in forma semplificata i materiali della tradizione manoscritta non assunti a testo, lezioni ritenute non originarie e tuttavia appartenenti a pieno titolo alla lingua dell'epoca, o perlomeno alla lingua del copista. Nella nostra prospettiva, documentando la fenomenologia testuale e linguistica della trasmissione manoscritta e il processo storico della copia, ci consegna una sorta di fotografia di una più complessa realtà lessicale, meritoriamente salvaguardata in una zona dell'edizione non sempre considerata come sarebbe conveniente. Per quanto possibile, tale realtà viene trasferita nel VD. A questa prima decisiva selezione dei materiali di base si aggiungono altri, a partire da quelli derivanti dallo spoglio della "storica" edizione della Crusca del 1595, edizione "criticamente" impostata sul confronto di un centinaio di manoscritti e stampe che ancora oggi, a distanza di secoli, si fa ammirare per la lungimirante intenzione degli Accademici di fornire ai redattori del grande *Vocabolario* della lingua italiana (che sarebbe apparso nel 1612) un testo della *Commedia* filologicamente fondato. Inoltre nel nostro lavoro si assumono anche i dati provenienti dallo spoglio sistematico delle edizioni successive a quella di Petrocchi (Lanza, Sanguineti, Bellomo, Inglese) e da lavori testualmente rilevanti, in grado di fornire elementi utili, in primo luogo le ricerche di Trovato e del suo gruppo di collaboratori. Operativamente vengono sfruttate le *Opere di Dante lemmatizzate (DanteSearch)*, progetto coordinato da M. Tavoni, e le *Concordanze delle Opere volgari e delle Opere latine di Dante Alighieri*, contenute in un CD-rom allegato alla riedizione delle *Opere* di Dante della Società Dantesca Italiana, riveduta da D. De Robertis e G. Breschi. E inoltre, come è ovvio, tutto quanto di utile e di sfruttabile proviene da banche-dati, archivi, corpora, commenti, enciclopedie, vocabolari, lessici, studi specifici vari. **Impatto scientifico.** Un *Vocabolario* delle opere di Dante, realizzato con i criteri sopra descritti, costituisce una assoluta novità nel panorama della pur ricchissima bibliografia dedicata al poeta, nella quale manca uno strumento che illustri sistematicamente e integralmente il lessico dantesco, interpretandolo alla luce delle più moderne acquisizioni e dei più recenti strumenti di analisi. Le Schede del VD, che integrano l'aspetto tradizionale dei lessici con un'interfaccia dinamica e modulare, e i sistemi di ricerca previsti, che consentono interrogazioni semplici o molto avanzate, predispongono lo strumento a tipi di fruizione differenziata, assecondando sia le esigenze del lettore esperto interessato all'indagine linguistica e filologica, sia le esigenze di un più vasto pubblico, italiano e straniero. **Obiettivi di questa proposta.** Dopo numerosi interventi in cui sono stati indicati i presupposti metodologici e i lavori in corso del VD (cfr. bibliografia), una prima presentazione ufficiale si terrà il 1° ottobre 2018 a Firenze, presso la Villa Medicea di Castello (sede dell'Accademia della Crusca), nel corso di una pubblica tornata accademica. Nell'intervento proposto per il Congresso di Copenhagen, il Gruppo di Redazione del VD, a partire dalla pubblicazione delle prime schede, farà

il punto sullo stato dei lavori e presenterà le innovazioni sostanziali, di metodo e di contenuto, che il VD può offrire alla lessicografia e alla storia della lingua.

CORBELLA, Dolores : *Diccionario de portuguesismos atlánticos* y estratigrafía del español americano.

Durante los últimos años hemos estado trabajando en la elaboración de una base de datos lexicográfica relacional de portuguesismos atlánticos en la que hemos recopilado toda la información metalexicográfica y los testimonios documentales que avalan la presencia e incorporación de los préstamos de esta procedencia en el llamado español meridional. Esta base de datos tiene como objetivo último la publicación de un diccionario de préstamos portugueses al español atlántico. En una primera etapa del proyecto, dedicada a los siglos iniciales de la expansión colonial, analizamos los diversos ámbitos temáticos y la diatopía donde la interferencia fue más intensa. Para ello investigamos, por un lado, si se puede constatar el registro de esas voces en tierras peninsulares (en las zonas fronterizas andaluzas y extremeñas) y comprobamos, por otra parte, la integración efectiva de esos préstamos en las nuevas tierras conquistadas *além-mar*, primero en el archipiélago canario y después en el Atlántico occidental, especialmente en el Caribe insular y continental, así como en Nueva España (Corbella/Fajardo 2017). Una característica general de todas las sociedades insulares renacentistas (entendidas como sociedades de frontera) fue la movilidad social, de ahí que las Canarias y después las Antillas fueran en los primeros momentos polos de atracción y, con posterioridad, áreas centrífugas, lo que provocaría una emigración interregional y, a la postre, contactos transoceánicos que quedaron perfectamente reflejados en la lengua. Ello dio lugar a la adopción de voces exclusivas y, a la vez, a la difusión de una buena parte de préstamos interregionales o interdialectales, tal como habían sugerido Salvador (1967) y Granda (1968) en sendos estudios clásicos sobre este tema. En la segunda fase del proyecto recopilamos el resto de lusismos atlánticos que se han integrado en español en época más reciente, especialmente a partir de los siglos XVII y XVIII, y que han tenido dos vías principales de entrada: el brasileño y el establecimiento de grupos poblacionales procedentes de Portugal en determinadas áreas. Hay que tener en cuenta, además, el importante aporte de los flujos migratorios gallegos en zonas como el Río de la Plata (también en Cuba y Venezuela), que ha influido, sin duda, en el arraigo de algunas voces del fondo patrimonial gallego-portugués o del gallego exclusivamente. El contacto con el brasileño ha dado origen a tres nuevos tipos de trasvases léxicos: los brasileñismos (o brasilerismos) propiamente dichos, esto es, palabras o acepciones propias del portugués de Brasil surgidas ya en contexto americano; los afrobrasilerismos, en los que la lengua portuguesa ha actuado como transmisora de ese léxico propio de los esclavos de procedencia africana (como, por ejemplo, *mucama*) y, por último, los indobrasileñismos (o tupibrasilerismos como *jacana*), esto es, las voces indígenas americanas que han tenido en el portugués su vía de entrada en español. A medida que avanzamos en el tiempo, se puede comprobar que las relaciones se intensificaron en la frontera brasileña y que las situaciones de contacto de lenguas produjeron un trasvase importante de léxico en las regiones limítrofes de Bolivia, Colombia, Argentina, Paraguay y Uruguay. Pero, del mismo modo, no fue menos importante la creación de nuevos asentamientos con colonos españoles traídos de la otra orilla del Atlántico, como sucedió en Luisiana, en el Río de la Plata (Montevideo) o en la Península de Yucatán (Chetumal, Campeche, Bacalar), que traían en su bagaje lingüístico voces de origen portugués que ya habían incorporado en sus modalidades dialectales. Esta base de datos que

contiene la historia atlántica de cada uno de los portuguesismos recopilados ofrece, además, la posibilidad de recuperar la información por los campos relacionales creados al efecto, como la tipología del préstamo, la categoría gramatical, la fecha de la primera documentación, la adscripción temática o su distribución geolectal. En esta comunicación nos centraremos precisamente en este último aspecto, ya que el programa permite visualizar los registros y mostrarlos de dos maneras diferentes: · Atendiendo al eje espacial únicamente, es posible mostrar la dispersión geográfica por países (y por regiones, dentro de cada nación) o por áreas dialectales, con el fin de conocer el paisaje lingüístico de este tipo de interferencia, si el préstamo constituye una de las características definidoras de una zona del país frente al conjunto, de una determinada variedad frente al resto o de un área dialectal concreta del español atlántico (andaluz, canario, Antillas, Caribe continental, Chile, Río de la Plata, Área andina, México y Centroamérica o Estados Unidos). · Combinando los ejes diatópico y diacrónico, el programa añade además, como novedad técnica, la posibilidad de representar cartográficamente la historia léxica de cada palabra (una especie de estratigrafía dialectal que representa, a un tiempo, la progresión cronológica y diatópica del término analizado, teniendo en cuenta la evolución semántica que ha experimentado). A modo de atlas lingüístico diacrónico, la memoria histórica de cada palabra queda perfectamente contextualizada, al menos para aquellas voces en las que la información acopiada resulta significativa, ya que la evolución del lema queda asociada a su representación cartográfica. Se trata, en definitiva, de un proyecto de lexicografía histórica en el que se combinan los registros textuales y los aspectos relativos a la cronología, dispersión y difusión topolectal de los préstamos, lo que en sí mismo constituye una novedad en la investigación diacrónica del léxico español, al poder confirmar con datos fidedignos extraídos del corpus las etimologías propuestas, sus más antiguas documentaciones y la ruta seguida por los términos analizados. Hemos tomado como referencia el modelo desarrollado por Perea (2012) para el catalán. En español, la técnica de cartografiado de los materiales diacrónicos solamente se está llevando a cabo en el *Atlas lingüístico diacrónico e interactivo de la Comunidad de Madrid (ALDICAM-CM)*, desarrollado por el grupo GITHE de la Universidad de Alcalá de Henares. En cuanto a la novedad, la elaboración de un *Diccionario de portuguesismos atlánticos (DPA)* surge de la necesidad de cubrir un vacío en la lexicografía española, con el fin de ofrecer datos contrastados sobre las etimologías lusas y corregir los numerosos errores que se han generalizado y perpetuado en los repertorios lexicográficos. El *DPA* nace, además, como no podía ser de otra manera en el siglo XXI, con vocación digital, como base de datos relacional pero también como material siempre abierto a nuevas incorporaciones, registros documentales o precisiones que vayan acopiándose en el futuro y que permitan su actualización.

DÍAZ RODRÍGUEZ, Cristian & PRIVAT, Maryse : Del tratamiento de la variación diatópica en los diccionarios bilingües francés-español: ¿qué francés y qué español?

Los currículos académicos que sientan las bases por las que se rige la enseñanza del español y del francés como lenguas extranjeras (en adelante, ELE y FLE) parecen coincidir en la necesidad de fomentar una política lingüística que, lejos de reducir la lengua y la cultura a un patrimonio exclusivo de España y Francia respectivamente, abogue por un tratamiento de corte panhispánico y panfrancófono de la lengua, que dé cuenta de la riqueza de ambas esferas lingüísticas y culturales. Baste citar, como prueba testimonial, la cantidad de documentos reales extraídos de la literatura, del arte o de las costumbres de países francófonos o latinoamericanos en los manuales de ELE o de FLE,

en los que abundan los referentes culturales endémicos y los elementos léxicos que reflejan la variación diatópica. Las obras lexicográficas, especialmente las de tipo bilingüe, en nuestro caso concreto francés-español, en tanto que herramientas fundamentales que garantizan los procesos de comprensión y producción de los aprendientes alófonos, no pueden ser ajenas a este enfoque de las lenguas y, por ello, deben incluir en sus repertorios aquellos lemas que permitan ofrecer un panorama fiel y representativo de la variación diatópica. No en vano, en la mayoría de las fichas técnicas de los diccionarios, el tratamiento de la variación diatópica aparece, aunque solo sea de manera cuantitativa, como uno de los principales descriptores de las obras lexicográficas. En esta comunicación, pretendemos realizar un estudio del tratamiento lexicográfico que recibe la variación diatópica del español y del francés en los diccionarios bilingües. Para tal fin, nos serviremos de los dos diccionarios bilingües con motor de búsqueda más representativos desde un punto de vista didáctico, a saber: el *Vox-Larousse* y el *Collins-Robert*. De hecho, para nuestro estudio, utilizaremos diferentes versiones de estos diccionarios, en función del usuario para el que están diseñados: Avanzado, Escolar, Esencial. Si bien es cierto que las características de nuestro objeto de estudio exigen un tratamiento cuantitativo del mismo, tanto más cuanto que el carácter contrastivo está triplemente presente—norma vs variedad diatópica; francés vs español; diferentes diccionarios de nuestro corpus—, privilegiaremos, no obstante, la descripción cualitativa del tratamiento que se hace de la diatopía. Esta decisión metodológica se justifica por el hecho de que, ante la dificultad de juzgar sobre la representatividad de los lemas diatópicos—lo que implicaría llevar a cabo estudios frecuenciales éticos y émicos— y, por tanto, sobre la pertinencia de su inclusión en los diccionarios bilingües, decidimos centrarnos en el análisis del tratamiento que recibe la información que ha decidido incluirse. En un primer momento, nos interesaremos por la descripción que se hace de la variedad diatópica en los prefacios de cada uno de los diccionarios que componen nuestro corpus, pues esta va a condicionar su alcance y las marcas diatópicas utilizadas. Así, se comprobará que, en el caso del francés, el pretendido tratamiento panfrancófono se restringe exclusivamente a la inclusión de términos procedentes de las variedades belga, suiza y canadiense, obviando gran parte, y no la menor, de la francofonía. En cuanto al tratamiento que recibe la diatopía española en los diccionarios estudiados, se observará que, tan solo en algunos casos, el léxico ajeno a la Península Ibérica se identifica con marcas o etiquetas que ubican, con mayor o menor precisión, su procedencia (los regionalismos carecen, generalmente, de marcación en estos diccionarios), v. g. *Arg.* (Argentina) vs *RPlat* (Río de la Plata). De hecho, la mayoría de las veces, el léxico representativo de las variedades diatópicas queda amalgamado bajo la etiqueta *Am.* (= americanismo), uniformizando, o más bien falsificando, una realidad léxica americana que está lejos de ser homogénea. Además de señalar las discordancias de tratamiento que recibe un mismo lema en los distintos diccionarios de nuestro corpus, también pondremos en evidencia la falta de sistematicidad, y, en algunos casos, de congruencia, que existe entre las diferentes versiones de un mismo diccionario (en papel, electrónico, en línea) o, incluso, dentro de un mismo diccionario. En aras de ofrecer un estudio completo sobre las marcas diatópicas, recurriremos igualmente a los diccionarios monolingües a partir de los cuales se han construido las entradas de nuestros diccionarios bilingües (principalmente el *Petit Robert*, el *Trésor de la Langue Française informatisé* y el *Dictionnaire Larousse* para el francés, y el *Diccionario de la Lengua Española* para el español), con la finalidad de ver si existen diferencias sustanciales en lo que a la marcación se refiere. En una segunda parte, analizaremos en qué medida los diccionarios bilingües cumplen con el principio de reversibilidad, que se les presupone, a la hora de tratar el léxico diatópico. De hecho, comprobaremos que, en la mayoría de los casos, los diccionarios bilingües solo contemplan la diatopía desde un enfoque decodificador, es decir, cuando el diccionario se utiliza como una

herramienta de comprensión. Sin embargo, adolecen de representatividad si se examinan desde el punto de vista de la codificación-producción. Así, por ejemplo, el usuario podrá encontrar fácilmente qué significan los términos diatópicos *nonante* (Bélgica y Suiza) y *mamadera* (América), pues ambos poseen una entrada propia en los diccionarios bilingües. Sin embargo, no se les ofrecerá dicha variante cuando se consulten las entradas de sus homólogos traductológicos *noventa*, en español, y *biberon*, en francés. Señalaremos, además, la deficiencia con la que se registra la diatopía semántica, a saber, cuando un significante compartido por toda la comunidad de hablantes desarrolla una acepción endémica de un geolecto específico (v. g. *arrugar* (Cuba) ‘fastidiar, molestar a alguien’; *sonner* (Bélgica) ‘llamar’), así como la diatopía fraseológica, ya sea formal, (v. g. *ser la guinda del pastel* vs *ser la frutilla de la torta* (Argentina); *en avoir ras la casquette* vs *en avoir plein son capot* (Quebec)), o semántica (v. g. *avoir du chien* ‘tener sexapil’ vs ‘ser astuto’ (Quebec)). Nuestro análisis nos llevará a concluir que, frente a la pretendida tendencia al panhispanismo o a la panfrancofonía, que preconizan todos los planes curriculares de ELE y FLE, las obras lexicográficas bilingües francés-español siguen manteniendo una actitud bastante eurocéntrica, e incluso nacional, que obedece por igual a políticas lingüísticas y políticas editoriales.

D’ONGHIA, Luca & TOMASIN, Lorenzo : Per un (meta)vocabolario storico ed etimologico del veneziano.

La comunicazione intende illustrare la struttura, le caratteristiche e i problemi di un metavocabolario storico ed etimologico del veneziano che si vorrebbe mettere presto in cantiere. La scelta del veneziano dipende in prima battuta dagli interessi di ricerca dei proponenti, oltre che dall’eccezionale continuità documentaria di cui gode quella varietà, diffusa e scritta a Venezia e in molti punti del Mediterraneo per buona parte dell’età moderna (esiste, come è noto, un veneziano ‘coloniale’ o « de là da mar », studiato tra i primi da Gianfranco Folena e diffuso tra l’altro a Ragusa, Zara, Creta, Cipro e così via, cfr. Baglioni in c.d.s.). Il veneziano è inoltre, nel panorama dei volgari (e poi dialetti) italiani, uno dei meglio studiati in molte sue fasi (dalla raccolta di testi antichi procurata da Stussi 1965 fino alla sistemazione complessiva offerta da Ferguson 2007), ma anche uno di quelli sondati più a fondo da una tradizione lessicografica ricchissima, che ha episodi significativi già tra Sei- e Settecento e che culmina con il Dizionario del veneziano di Giuseppe Boerio (Boerio 1829 e edizioni successive), ancor oggi indispensabile strumento di lavoro per chiunque studi cose lagunari. Non è tutto: il veneziano è anche una delle poche varietà a godere anche di una nutrita lessicografia storica (nel senso di retrospettiva e documentaria), dedicata a periodi o autori rilevanti. È questo il caso dei vocabolari rispettivamente dedicati al veneziano di Carlo Goldoni da Gianfranco Folena (Folena 1993) e al veneziano cinquecentesco da Manlio Cortelazzo (Cortelazzo 2007). Accanto a opere di larga mole come queste ultime ce ne sono poi numerose altre variamente rilevanti : basterà rammentare qui lavori diversissimi ma egualmente preziosi come le Etimologie venete di Angelico Prati (Prati 1968), gli studi di Elke Sallach sul lessico veneziano cinquecentesco (Sallach 1993) e la recente monografia di Francesco Crifò dedicata a Marin Sanudo (Crifò 2016). Da un lato si ha dunque una cospicua, preziosissima quantità di dati – di provenienza propriamente o latamente lessicografica (dizionari, studi dedicati al lessico e glossari di edizioni) – che sarebbe utile poter mettere a sistema disponendone per così dire sinotticamente e simultaneamente; d’altro canto i tempi e i costi di realizzazione di un vocabolario storico tradizionalmente concepito – fondato cioè sulla costruzione sorvegliata di un corpus adeguato, sulla sua lemmatizzazione e sulla confezione delle voci – non

appaiono realisticamente affrontabili nel quadro attuale della ricerca umanistica. Non sono mancati in effetti progetti assai ambiziosi, capaci di produrre eccellenti edizioni per singoli testi, ma purtroppo non arrivati all'esito lessicografico che si proponevano : è il caso, per l'ambito di cui si parla, dell'importante progetto di ricerca italiano « Vocabolario storico dei dialetti veneti », attivo in Italia tra il 1997 e il 2003. Data questa situazione, la storia (lessicale) del veneziano potrebbe essere efficacemente ricostruita puntando sull'allestimento di un metavocabolario, che avrebbe il vantaggio di non dover passare attraverso l'edizione (o riedizione) di una miriade di testi e quindi attraverso il loro trattamento informatico e la loro lemmatizzazione. Il metavocabolario metterebbe a frutto principalmente le fonti seguenti (ci limitiamo di norma a due o tre esempi per ogni tipo): - tradizione lessicografica veneziana : per esempio Patriarchi 1775 (e edd. seguenti), Boerio 1829 (e edd. seguenti), Ninni 1890; - tradizione lessicografica veneziana monografica o settoriale : per esempio Folena 1993, Moretti 2002, Cortelazzo 2007. - tradizione 'paralessicografica' veneziana : per esempio Varotari 1671 e Muazzo 2008 (che andranno sottoposti per porzioni significative a spoglio manuale); - tradizione lessicografica non veneziana dalla quale sia possibile estrarre materiali veneziani : per esempio Paccagnella 2012, il Tesoro della Lingua Italiana delle Origini (che darà un contributo decisivo alla copertura del periodo fino al 1375), il Dizionario Etimologico della Lingua Italiana, gli indici dell'Atlante linguistico Italo-Svizzero; - lessicografica etimologica: per esempio Prati 1968 e i materiali veneziani estraibili dal Lessico Etimologico Italiano e dal DELIN di Cortelazzo-Zolli; - glossari o note lessicali poste a corredo di edizioni critiche affidabili : per esempio i glossari in calce a Stussi 1965, o le note lessicali che accompagnano le edizioni in Tomasin 2012, Tomasin 2013 e nelle numerose altre edizioni recenti di testi veneziani soprattutto medievali; - studi dedicati al patrimonio lessicale o alla storia linguistica di Venezia : per esempio Sallach 1993, Tomasin 2001, D'Onghia 2009, D'Onghia 2011, Ferguson 2012, Crifò 2016. Sfruttando sistematicamente questo patrimonio si potrà ottenere, per ogni parola, una documentazione diacronica ricca e a maglie relativamente strette : per la fase medievale, oltre alla bibliografia specialistica, si può infatti contare sul Tesoro della Lingua Italiana delle Origini e sulla sua banca-dati ; per la fase rinascimentale su Cortelazzo 2007 ; per la fase settecentesca su Folena 1993 e Patriarchi 1775 ; per la fase ottocentesca su Boerio 1829 ; per la fase otto- e novecentesca su un cospicuo insieme di più recenti vocabolari (per es. Zambon 2009). Per alcuni periodi attualmente meno studiati, come il XIV e il XVII secolo, sarà opportuno avviare l'allestimento di glossari compilati su opere particolarmente significative : alcune tesi di dottorato in corso stanno già sperimentando la realizzazione di glossari mirati dedicati a singole opere o a piccoli corpora relativi a periodi già ricchi di materiali édit, ma poveri di trattazione lessicografica (ad esempio il Quattrocento con i suoi numerosi testi tecnici marineschi, o il Seicento con i rifacimenti dialettali di capolavori letterari del Rinascimento: cfr. rispettivamente Conterio 1991, Pittarello 2006, e Mondini 2002). Ogni voce del metavocabolario – di cui si intende offrire qualche saggio nella comunicazione – avrà una struttura in cui saranno senz'altro presenti : (1) regesto, in ordine cronologico, della lessicografia e dei glossari esistenti; (2) scheda etimologica; (3) scheda bibliografica. La pubblicazione delle voci avverrà in linea, sfruttando una piattaforma che dovrà essere interoperabile con quelle delle maggiori imprese analoghe, segnatamente il Tesoro della Lingua Italiana delle Origini.

FAJARDO AGUIRRE, Alejandro : Portuguesismos en la lexicografía española: incorporación, tipología y áreas temáticas.

La incorporación de léxico de origen portugués o transmitido por mediación de esta lengua ha tenido una importante repercusión en la configuración del vocabulario español, aunque, en contraste con los elementos de otros orígenes, no siempre haya merecido la atención de los especialistas en la justa medida. Algunos destacados lingüistas han llamado la atención sobre este hecho, p. ej., Salvador (1966), Pérez Vidal (1991), de Granda (1968), Frago (1999), Lüdkte (2014), entre otros. Trabajos más recientes (Corbella / Fajardo 2017) han permitido obtener una amplia visión de la importancia de los contactos léxicos entre el español y el portugués. Sin embargo, el registro de los préstamos portugueses en la lexicografía española ha sido desigual y, en general, insuficiente. Esto se explica, en parte, por su diversidad tipológica y por la especialización temática de una parte de ellos. Pero, sobre todo, hay que tener en cuenta que su adaptación al español se fue produciendo de una forma natural, con frecuentes bifurcaciones, palabras cognadas y calcos. Esto provoca dificultades en su identificación y en su registro y marcación lexicográfica. Se suma, además, la circunstancia de que hay contactos poco documentados (p. ej., en los núcleos fronterizos en América) y que la lexicografía dialectal de muchas áreas fronterizas con Brasil necesita aún muchas investigaciones. Este conjunto de factores dificulta, sin duda, tanto el registro lexicográfico de los portuguesismos como su adecuada marcación. Desde un punto de vista cronológico, la llegada y adopción de portuguesismos se produjo en diversas oleadas: las primeras aportaciones medievales corresponden a elementos cuyo origen se documenta en gallego-portugués y que se extendieron por el ámbito territorial iberorrománico antes de la expansión atlántica de los reinos de Castilla y Portugal. En portugués clásico, se dio un fuerte bilingüismo propiciado por la unión de coronas, periodo que coincide con las fechas de primera aparición en los textos de una gran parte de los préstamos; a ellos se suman los exotismos de orígenes lejanos, muchos procedentes de las Indias Orientales, transmitidos al español por intermediación de la lengua vecina, cuyos hablantes, ya fueran navegantes, comerciantes o misioneros, fueron pioneros en gran parte de aquellas regiones lejanas; es, también, la época de los portuguesismos atlánticos, que se propagan por el océano hacia América a través de los emigrantes. En las Islas Canarias, donde la presencia de portugueses fue de gran importancia en distintas actividades, la influencia fue muy importante. También en las zonas fronterizas de América en las que ambas lenguas han estado en permanente contacto lingüístico, el intercambio de elementos léxicos en las dos direcciones ha sido continuo en los últimos siglos. Su asimilación es extraordinariamente rápida y la fusión da lugar a una gran productividad en la creación de palabras ya plenamente españolas, lo que es otra dificultad para la identificación del préstamo. Veremos en qué medida, la lexicografía del español ha estado más o menos abierta a la recepción de los portuguesismos de distintos orígenes. La compilación de portuguesismos a partir de su registro en repertorios dialectales se ha llevado a cabo con mayor profundidad en determinadas zonas, p. ej., en Canarias, debido a la importancia de este elemento entre los pobladores de las islas. Pérez-Vidal (1991), reunió más de 900 de estos préstamos, que se agrupan temáticamente en 26 campos. Muchos de ellos son frecuentes en el habla coloquial, aunque el hablante ignore su origen y los tome, a veces, por palabras que identifica con lo regional; muchas otras voces de este origen, sin embargo, han quedado en desuso. Si extraemos los datos del DECH de Corominas-Pascual (1991), comprobamos que se registran en el 150 entradas de posible o seguro origen portugués. Un análisis por campos nos permite analizar su reparto. Como veremos, gran parte de las voces registradas en este diccionario corresponden al léxico relacionado con el mar, la flora es el segundo grupo en número de voces y las denominaciones de la fauna son también frecuentes. A

partir de estos datos, realizamos una comparación con otras fuentes lexicográficas de interés: en primer lugar, se lleva a cabo el contraste de estos datos con los registros de la última edición del DLE (2014) y con otros estudios sobre campos específicos que han aportado datos inéditos. Podremos así extraer conclusiones sobre el registro de las voces y la adecuación de la marcación que han llevado a cabo los diccionarios del español, así como la mayor o menor atención prestada a campos determinados campos temáticos. En la comunicación, se presentarán, finalmente, las más recientes informaciones cuantitativas y cualitativas resultantes del análisis de los registros de la base de datos del proyecto “Portuguesismos Atlánticos (ss. XVIII-XX) FFI2016-76154-P”

GIMÉNEZ EGUÍBAR, Patricia & SÁNCHEZ GONZÁLEZ DE HERRERO, M^a Nieves : Estudio de los nombres de oficios y cargos en la documentación medieval del monasterio de San Andrés de Vega de Espinareda (León) (Siglo XIII).

Un hecho unánimemente admitido en los estudios dialectales de la Península Ibérica es la imposibilidad de establecer fronteras lingüísticas nítidas entre las variedades norteñas surgidas del latín en la Edad Media y la existencia de un continuo lingüístico. En la época medieval hemos de considerar además, entre otros hechos, la movilidad de las fronteras de los distintos reinos y la falta de homogeneidad lingüística en su interior. En palabras de Moreno Fernández (2005: 82): “Cada territorio reunía elementos diferentes que se conjugaban de modo diverso para la creación de los nuevos instrumentos lingüísticos, por más que muchos de esos elementos fueran compartidos”. Si atendemos a la consideración de variedades bajomedievales escritas próximas entre sí, y más concretamente a los límites escrito-lingüísticos de dos variedades vecinas como la galaica y la leonesa, vemos que en estudios previos se ha puesto de manifiesto la problemática delimitación entre áreas de Galicia limítrofes con las que continúan en las actuales provincias de León y Zamora. Souto Cabo (2008: 171), por ejemplo, afirma que el debate puede surgir en primer lugar a la hora de definir, desde un punto de vista idiomático, la producción emanada de los monasterios de Castañeda (Calende), Carracedo (Carracedelo) y San Pedro de Montes (Ponferrada). Es precisamente en una de estas zonas de continuidad y contacto entre el gallego y el leonés, en el Bierzo, donde vamos a situar el análisis léxico propuesto, tomando como objeto de estudio la Colección Documental del Monasterio de San Andrés de Espinareda (1043-1428) (Jiménez Suárez 2005). Consideramos que dicha edición, así como la colección en la que se inserta, reúne la fiabilidad y calidad suficientes como para ser objeto de un estudio lingüístico; no obstante, no descartamos la posibilidad de acudir a los originales, lo que hasta ahora no ha sido posible. El trabajo mencionado acoge la edición íntegra de los documentos conservados, en original o copia, hasta el año 1300; del resto, hasta 1428, consta únicamente el regesto. Son 73 testimonios hasta 1298, sumando originales y copias, con inclusión de documentos de cancillería. Entre los 73 hay 20 copias, algunas de cancillería real, copias que eliminamos en su totalidad, porque desde el punto de vista filológico interesan los originales, de manera que nos quedamos con 53. De 1043 a 1244 contamos con 14 documentos de los cuales 4 son copias mientras que los otros diez, originales, están redactados siguiendo la norma latina medieval; es decir, vamos a basar nuestro estudio en 43 documentos originales datados entre 1253 y 1294 y escritos en romance, o, al menos, con partes de la disposición en romance. La riqueza que ofrecen estos testimonios para la historia del contacto entre las dos variedades nos atrajo ya en una primera lectura y decidimos llevar a cabo un estudio detallado. El resultado de este primer acercamiento lo presentaremos próximamente en el North American Symposium of Galician Studies, que se celebrará

en Denver (Colorado) entre el 18 y el 20 de octubre de 2018. Las ponencias, ya aceptadas tras su revisión y anunciadas, llevan por título: “Estudio de los rasgos de continuidad del gallego en documentos del monasterio de San Andrés de Espinareda (León) en el siglo XIII: niveles graficofonético y morfosintáctico” (Sánchez González de Herrero) y “nivel léxico” (Giménez Eguíbar). En este primer acercamiento de carácter general al léxico, observamos el interés de los nombres de cargos y oficios que frecuentemente, aunque no de manera exclusiva, acompañan a las listas de testigos del acto jurídico, siguiendo la costumbre que va haciéndose cada vez más habitual a partir del siglo XI en la documentación peninsular, bien precediendo al nombre propio, bien como designación complementaria tras él (Clavería Nadal 2007: 538). Entre ellos encontramos algún posible diatopismo, del tipo clerizón (DCECH, s.v. clero), aunque no es lo más frecuente. Un aspecto importante es el que nos ofrece el sistema de derivación y composición que siguen, mercador, hostalarío, tendeyro o zapatero. En este marco, muestra especial significación la alternancia ei/e, procedente de AJ, como era esperable (Morala 2004: 557), presente en voces del tipo meyrino/merino, o en los abundantes derivados de -ARIUS, del tipo uestiareyro, raçoneyro, palleyro, sarralleyro, etc., que conviven con carpintero, carnicero, colchero o escudero, etc., en ocasiones en un mismo testimonio (documento 67 de la colección citada, por ejemplo). Pretendemos analizar con detalle esta alternancia, de interés en la dialectología histórica, para tratar de descubrir sus posibles razones (relación con determinados notarios, lugar de emisión o parte del documento en que se inserta, entre otros) con la finalidad de aclarar, si es posible, la disposición espacial de ciertos hechos lingüísticos. No buscamos, en definitiva, la presencia de diatopismos sino la distribución de voces y fenómenos concretos en el marco de una comunidad, galaico-leonesa en este caso, al margen de los límites de las lenguas estudiadas, así como las causas que podrían dar cuenta de dicha distribución.

GLESSGEN, Martin : Un *Nouveau Godefroy* : la description lexicographique et l'apport lexicologique des textes documentaires français médiévaux.

En 1998, nous avons entrepris une réflexion systématique autour des textes documentaires français médiévaux (Glessgen 2001), réflexion qui a débouché relativement vite à l'idée de concevoir un ‘Nouveau Godefroy’ (Glessgen 2003). Notre idée, aussi simple qu'exigeante, était celle (1) de réunir en une base de données relationnelle l'édition d'un nombre important de textes documentaires pour (2) permettre ensuite la rédaction d'un dictionnaire thématique, ciblé sur un ensemble de traditions de discours par trop négligé en linguistique diachronique. Pour le premier volet de notre entreprise, nous avons constitué à travers les quinze dernières années la base textuelle des Documents linguistiques galloromans, soutenue par l'École Nationale des Chartes et aidée par de nombreux collègues, amis et élèves. Actuellement, cette base comporte 1 250 000 occurrences en près de 3 000 actes provenant de textes en territoire d'oïl ainsi que 7 000 000 occurrences dans des textes provenant du territoire francoprovençal et étant partiellement rédigés en français (DocLing en ligne, cf. Glessgen 2015 et 2018). Pour le volet lexicographique, nous avons élaboré, également à travers plusieurs années et avec un soutien informatique professionnel, un logiciel (Phoenix) pour différents types d'interrogation et, également, pour la lemmatisation des formes, qui est désormais achevée. Pour l'exploitation lexicographique de ces données nouvelles, nous avons toutefois renoncé à la constitution d'une nouvelle base lexicographique à part. L'idée que nous avons retenue dès 2013 a été d'intégrer notre vocabulaire dans la base de données du Dictionnaire étymologique de l'ancien français, DEAFpré (projet présenté lors du Congrès de Nancy, cf. Glessgen/Tittel en prép.). La programmation du

module d'intégration – très exigeante – étant achevée à la fin de l'année 2017, nous avons pu transposer avec notre collaboratrice Marguerite Dallas le vocabulaire des actes au sein de cette plateforme. Actuellement, nous préparons la rédaction des premiers articles dans le cadre du DEAF, ce qui est une expérience novatrice en lexicographie électronique, ce dictionnaire de référence s'ouvrant ainsi à des intégrations extérieures. Nous sommes à présent au seuil de la rédaction du 'Nouveau Godefroy', objectif que nous avons poursuivi dès le début. À partir d'une brève présentation des différentes étapes et volets du travail achevés et en cours, nous souhaiterions entreprendre une réflexion plus approfondie et plus globale sur les apports de ce projet lexicographique, qui ne se substituera nullement au vénérable dictionnaire de Frédéric Godefroy. Ses dix volumes englobent en effet tous les genres textuels qu'ils soient littéraires, documentaires ou, dans une moindre mesure, scientifiques. Pour les parts littéraires et scientifiques, le Godefroy a désormais été remplacé par les grands dictionnaires de l'ancienne langue, le Tobler-Lommatzsch, le DEAF, le DMF et l'AND. L'apport de notre projet se concentre toutefois sur les textes documentaires, très présents dans Godefroy, mais relativement peu pris en considération par les autres dictionnaires. C'est ici que se place donc la particularité et le potentiel novateur du 'Nouveau Godefroy'. Nous essaierons de quantifier cet apport et de le mettre en perspective par rapport aux contenus du Godefroy et des autres dictionnaires médiévistes (époques, champs sémantiques, régionalité etc.). Enfin, nous mettrons en avant les nouvelles issues interprétatives qui découlent des innovations méthodologiques d'une lexicographie pleinement informatisée.

GONÇALVES, Maria Filomena : A REPRESENTAÇÃO DO PORTUGUÊS DO BRASIL EM DICIONÁRIOS DO SÉCULO XIX: LEXICOGRAFIA E IDEOLOGIA.

A comunicação que propomos tem como objectivo analisar o modo como, a partir de meados do século XIX, os dicionários publicados no Brasil apontam e reivindicam a especificidade do Português Brasileiro, em contraponto ao Português Europeu. O valor dessas obras é tanto maior quanto se sabe que a produção lexicográfica brasileira é tardia, situação que se deve ao facto de a impressão de livros em território brasileiro não ser permitida até às primeiras décadas do século XIX. Por isso, os dicionários e outras obras didácticas eram impressos exclusivamente em Portugal. As lutas liberais, por um lado, e, por outro, a independência do país em 1822 possibilitaram não só a difusão da imprensa como também a publicação de dicionários que, centrados no léxico brasileiro (os chamados "brasileirismos"), inauguraram a produção local (Cunha 1987; Oliveira 1998; Isquierdo 2006, 2007). Um ano antes da independência, saíra no Rio de Janeiro o Dicionario cardundatico ou explicação das phrazes dos carcundas (Lima 1821). Imitando as características de um dicionário, a obra tem intenção panfletária e arrola apenas palavras e expressões usadas no contexto da guerra ideológico-política entre liberais e antiliberais. Embora a consciência da identidade linguística brasileira fosse obviamente muito anterior à independência, é depois desta que a afirmação da variedade americana do Português se traduz em obras que expõem e sancionam as diferenças entre a norma brasileira e a norma europeia da Língua Portuguesa. Em 1826, no Atlas ethnographique du globe, ou classification des peuples anciens et modernes d'après leur langue, editado por Adrien Albi, o Visconde da Pedra Branca dera já a conhecer uma lista de vocábulos próprios do Português falado no Brasil, e, portanto, desconhecidos dos falantes europeus, mas a amostra apresentada pelo Visconde era apenas a assunção das diferenças entre duas variedades e respectivas normas. Em 1832 vem a lume o Diccionario da Lingua Brasileira, de Luís Maria da Silva Pinto, obra que, mau grado o título, não compila

propriamente o léxico brasileiro e também não é um dicionário original, lavrado inteiramente pelo seu autor e impressor, conforme mostrou Beserra Araujo (2009). É o primeiro dicionário impresso no Brasil e também a primeira a aludir à “língua brasileira”, em sintonia com a ideologia nacionalista do período pós-independência (Gonçalves 2012). Porém, este dicionário não inaugura verdadeiramente a lexicografia brasileira, revestindo-se de um valor mais simbólico do que linguístico. Só a partir de 1853 – data do Vocabulário Brasileiro, publicado por Brás da Costa Rubim – aparecem dicionários cujo objectivo é arrolar brasileirismos e oferecer o significado respectivo. Algumas das obras que vêm a lume não são, contudo, dicionários gerais, mas antes dicionários de regionalismos, como por exemplo os de Coruja (1852) ou Correa (1898), ambos centrados no léxico do Rio Grande do Sul, cuja variedade regional do Português (riograndense) se caracteriza uma notória influência lexical do espanhol falado na vizinha região platina (Gonçalves 2017). Tal como nesta, também em outras regiões brasileiras se fizeram compilações do léxico local que contribuíram quer para a lexicografia, quer para a dialectologia histórica da Língua Portuguesa no Brasil. Para o objectivo desta comunicação interessam-nos os dicionários que reúnem quer os chamados brasileirismos gerais (formas conhecidas e usadas por todos os falantes), quer vocábulos próprios das várias regiões brasileiras e que, em alguns verbetes, confrontam os usos brasileiros com os usos portugueses. Atentamos em duas obras publicadas produzidas em datas muitas próximas, a saber, o Dicionário Brasileiro da Língua Portuguesa (1875-1888), de Joaquim Macedo Soares (1838- 1905), e o Dicionário de Vocabulos Brasileiros (1889), do Visconde de BeaurepaireRohan (1812-1894). A selecção destas obras justifica-se pelo facto de terem em comum quatro aspectos: apresentarem uma lista de fontes, pretenderem inventariar o léxico geral brasileiro, incluírem a marcação lexicográfica da variação dialectal brasileira, estabelecerem, para alguns vocábulos, o confronto entre a acepção brasileira e a portuguesa. Quanto às fontes de uma e de outra, é de salientar que o Dicionário Brasileiro da Língua Portuguesa, cuja publicação completa não ocorreu em vida do autor, leva em consideração o conhecimento filológico/linguístico da segunda metade do século XIX, conforme denota a inclusão de nomes como Du Cange (ed. 1883-1887) e Friedrich Diez. Por outro lado, A partir dos dados analíticos, trata-se de identificar a clivagem linguística e ideológica entre o Português do Brasil e o Português Europeu, e, ao mesmo tempo, atestar a relação entre lexicografia e ideologia no Brasil de finais do século XIX (Coelho 2012). A metodologia consistirá na constituição de uma amostra extraída de ambos os dicionários por forma a comparar os seus enunciados lexicográficos, a estrutura destes, o sistema de marcação lexicográfica (variações e usos marcados), e bem assim o modo como o discurso de um e de outro dicionário constrói uma “representação” do léxico brasileiro e, por essa via, uma representação da variedade brasileira do Português e da sua “autonomia linguística” (Soares 1888: XVII). Esta proposta inscreve-se claramente nos objectivos da Secção 6 (Lexicografia) do XXIXe Congrès International de Linguistique et Philologie Romane porque trata de um tema relevante para a lexicografia das línguas românicas e, em concreto, por trazer contribuições para o estudo do léxico do Português como língua pluricêntrica, incluindo aspectos relativos ao contacto entre Português e Espanhol num território da chamada România Nova. Além disso, é de realçar que a dicionarística brasileira, que tem sido estudada sobretudo na perspectiva da análise do discurso, não tem sido alvo de estudo sistemático nos domínios da lexicologia e lexicografia históricas.

HEINZ, Matthias : Lessico italiano e contatto linguistico nella Romània: nuovi dati sui prestiti dall'italiano in spagnolo, portoghese e catalano. (poster)

La comunicazione/il poster si propone di presentare nuovi dati sulla presenza dei prestiti dall'italiano (italianismi) nelle lingue iberoromanze, basandosi sulla documentazione ormai disponibile on line nella banca dati lessicale OIM (Osservatorio degli italianismi nel mondo, progetto collaborativo internazionale dell'Accademia della Crusca, il cui nucleo risale al Dizionario di italianismi in francese, inglese, tedesco (DIFIT). Nell'ambito dei lavori di estensione delle raccolte di italianismi si hanno, oltre il corpus pubblicato nel 2013/2014, il quale raccoglie italianismi storici e contemporanei in francese, inglese e tedesco, le aggiunte recenti del polacco e dell'ungherese e raccolte parallele per lo spagnolo, il portoghese e il catalano. Queste ultime permettono di dare uno sguardo comparativo sulla penetrazione di italianismi nelle tre aree linguistiche attigue, contrassegnate da tendenze in parte convergenti, in parte individuali in quanto all'integrazione di materiale lessicale di origine italiana. Dalla prospettiva lessicografica e lessicologica la presentazione metterà a fuoco le sfide metodologiche per l'inserimento parallelo di una tale serie di raccolte lessicali nonché i potenziali della consultazione comparativa del corpus che ne risultano.

HUMMEL, Martin & GAZDIK, Anna : Le Dictionnaire historique de l'adjectif-adverbe.

Dans les grammaires traditionnelles, normatives et scolaires, les syntagmes du type *parler haut* apparaissent généralement comme un petit groupe d'expressions qui sont entrées dans l'usage comme exceptions face aux adverbes en *-ment* (cf., par exemple, Wagner/Pinchon 1993 : 150). Cependant, les grammaires scientifiques en soulignent la productivité en français moderne (Riegel *et al.* 2011 : 657-658, Wilmet 2010 : 423-425, Grevisse / Goosse 2008 : §§557, 963, 993, 994). Certains linguistes vont même plus loin et considèrent l'adverbe en *-ment* comme une forme traditionnelle mise en péril par la tendance du langage jeune à le tronquer improprement (Moignet 1981, vol. 1 : 52). Confrontés à des énoncés comme *J'y vais direct* ou *Elle joue un peu facile* [au tennis], les locuteurs ont tendance à partager ce point de vue (cf Grevisse / Goosse 2008 : §963). Or, sur le plan morphologique, l'apocope de *parler hautement* donnerait **parler haute* et non pas *parler haut*. L'usage systématique du masculin singulier pour l'adjectif en tant que modificateur adverbial contredit donc clairement cette vision des choses. Du point de vue historique, l'emploi adverbial de l'adjectif a une longue tradition gèneine qui ne dépend pas de l'apocope d'un suffixe et qui remonte au latin, voire à l'indo-européen. Les études plus fouillées de Heise (1912) et de Grundt (1972) mettent en évidence la longue tradition des adjectifs-adverbes du français. C'est ainsi que la place des adjectifs-adverbes dans la langue française doit être située entre tradition et innovation (cf. Goes 1999: 221). Situé dans un cadre de recherche comprenant plusieurs langues romanes, voire l'anglais, (Hummel 2013, 2014a,b, 2015, 2017a,b,c, 2018a,b,, sous presse a,b, Hummel & Kortschak 2018, Hummel & Kröll 2015, Hummel & Valera 2017), le *Dictionnaire historique de l'adjectif-adverbe* (DHAA) a pour objectif de doter la recherche de la base empirique nécessaire pour déterminer le rôle historique de l'adjectif-adverbe français La lexicalisation de certaines de ces expressions justifie l'idée d'un dictionnaire. Cependant, la productivité de l'adjectif-adverbe nous incite à combiner le dictionnaire avec le concept d'un trésor qui offre une documentation complète des phénomènes. En tout cas, qu'on regarde les expressions figées ou l'emploi productif, on ne saurait contredire Lucien Tesnière (1966 : 469) qui a qualifié ces constructions de "fort savoureuses". Il suffit de citer quelques exemples de notre banque de données, en ligne depuis 2005 (Hummel & Stiegler

2005) : Vil acheter et vendre chier (12^e) Manger froid, boire chaud, dormir couché, debout ; Un garçon com moi s'accommode de tout (17^e) Mais qui pense pis pense souvent juste (19^e) Visons haut, pensons vrai, marchons droit (19^e) J'achète frais, j'achète appétissant, j'achète sain parce que j'achète sous cellophane. (20^e) Chiez dur ou chiez mou, mais chiez dans le trou. (20^e) Le domaine de l'emploi adverbial de l'adjectif dépasse le syntagme verbal auquel se limite le DHAA. Le domaine d'usage comprend également la modification de l'adjectif (*fort important, étroit cousu*) et de l'adverbe (*fort bien*), ainsi que des fonctions d'ordre énonciatif (*Bref, il ne viendra pas*), y compris quelques marqueurs du discours (*Bon, c'est pas grave*), et ceci dans toutes les langues romanes. L'accord occasionnel des adjectifs-adverbes avec le sujet ou le complément d'objet du verbe (*Les larmes lui coulent faciles* (Meyer-Lübke 1974: 448)) ou avec l'adjectif modifié (*fleurs fraîches écloses, portes grandes ouvertes, toute grande*) met en évidence le rapport étroit avec la fonction adjectivale. Du point de vue sémantique, on peut distinguer la modification adverbiale de manière proprement dite (*aller direct, truander petit, aimer haut, refuser net*) de la modification de ce que l'on appelle traditionnellement *objet interne* du verbe (cf., par exemple, Moignet 1963: 178; Regula 1957 ; cf. Abeillé & Godard 2004, Abeillé & Mouret 2010, Noailly 1994, 1997). En effet, *voir grand, manger chaud*, etc. se réfèrent, logiquement, à une qualité de ce que l'on voit ou mange, ce qui les rapproche de la prédication seconde. De façon secondaire, l'emploi métaphorique *voir grand* 'être ambitieux' ou simplement abstrait *manger chaud* comme comportement culturel (*Les Français mangent chaud le soir*) brouillent les limites entre modification d'un objet interne et du verbe lui-même. Dans des cas comme *couper court*, on trouve parfois la flexion dans l'emploi concret (*Il les a coupés courts*), mais non pas dans l'emploi métaphorique (*Il coupe court* 'il interrompt'). Le DHAA lemmatise les syntagmes verbaux qui intègrent un adjectif-adverbe. La macrostructure alphabétique va donc de *abaisser bas* à *voter utile*. La banque de données contient environ 13 000 exemples attestés dans FRANTEXT et des corpus spécialisés portant sur l'Ancien et le Moyen Français. Le nombre d'adjectifs-adverbes différents tourne autour de 300, et celui des groupes syntagmatiques s'élève à 2000 environ, dont le dictionnaire reprendra une sélection. Ces données montrent que la représentation des adjectifs-adverbes dans les grammaires comme petite liste d'exceptions ne correspond pas à la réalité. Afin de pallier aux tendances « bon usage » de Frantext, le dictionnaire comprend également 3500 citations provenant de sources Internet pour documenter l'usage informel. Citons-en un exemple qui respecte l'orthographe originale : Bonjour Denis, ces voiles je les aient tes té avec du vent a plus de 60km/h en les laissant genre drapeau c' est te dire que ca les secouent, pour le moment ça fait deux an que je m'en sert et en transport je ne les menagent pas ca fait partie du test! je les *roulent* mais jamais *pliée*! [...] Re Denis, je te met une photo pour que tu te rende compte que je les *roulent petit* fait pas attention au bazar!!!! La communication présente le DHAA dont la publication est prévue pour 2019. Elle présentera également des résultats de recherche portant sur deux aspects lexicologiques : - L'appartenance des adjectifs-adverbes au lexique fondamental du français. - La structuration paradigmatique des adjectifs-adverbes (cf. *cravater large / lâche / étroit*).

KREMER, Sarah : Les nouveaux outils typographiques du FEW. (poster)

Le Französisches Etymologisches Wörterbuch (FEW), dictionnaire étymologique et historique du français, des dialectes d'oïl, du francoprovençal, de l'occitan et du gascon, a été initié par Walther von Wartburg en 1922. Il compte aujourd'hui 25 volumes, soit l'équivalent de plus de 16000 pages

d'articles. Depuis 2003, les nouveaux articles sont publiés sous la forme de documents numériques téléchargeables au format pdf sur le site de l'ATILF. L'intégralité du dictionnaire est également accessible en ligne et son indexation partielle simplifie sa consultation. Cette diffusion répond en partie aux difficultés d'accessibilité que pose le FEW, tant du point de vue matériel — le dictionnaire n'était disponible jusque-là qu'en bibliothèques spécialisées — que lexicographique — le discours lexicographique s'avère particulièrement dense et crypté. La dématérialisation du FEW se poursuit dans le cadre de son informatisation (Renders 2015). Trois volumes du dictionnaire sont actuellement traités par des algorithmes de rétroconversion afin de produire des articles au format XML, entièrement balisés, indexés et interrogeables, librement disponibles en ligne. Mais l'exploitation de ces données se heurte à des problèmes liés à l'affichage des contenus lexicographiques. Jusque-là, aucune police de caractères ne dispose de l'ensemble des signes typographiques nécessaires à la composition de la transcription phonétique mise en place par Wartburg. Au-delà de la seule disponibilité de ces caractères, la question de leur encodage est essentielle puisque seule une stratégie d'encodage pérenne peut garantir l'interopérabilité des données diffusées en ligne. Notre travail de thèse, mené dans le cadre d'une collaboration entre le laboratoire de sciences du langage Analyse et traitement informatique de la langue française (ATILF/CNRS) et l'Atelier national de recherche typographique (ANRT/ENSAD Nancy), a consisté à élaborer différents outils typographiques pour le FEW informatisé. Ce contexte de travail inédit a amené des linguistes, des informaticiens et des designers à produire des outils adaptés à l'affichage, la diffusion, la rédaction et la consultation du FEW. Une famille de caractères typographiques inédite a été spécialement créée pour les besoins du dictionnaire. Elle intègre l'ensemble des signes, caractères spéciaux et variantes stylistiques utilisés pour composer et formater les articles et propose une solution d'encodage pérenne pour l'ensemble des caractères multi-accentués utilisés par le FEW pour transcrire des relevés phonétiques. Le schéma d'encodage de ces polices s'appuie sur la récente intégration dans les plages du système d'encodage universel Unicode d'une série de caractères employés en dialectologie allemande et romane (The Unicode Consortium, 2014). Les polices développées pour le FEW sont parmi les premières à exploiter ces nouveaux points d'encodage. En présentant la méthodologie adoptée lors du développement des polices et des interfaces de consultation et de rédaction ainsi que leur mise en action, notre communication abordera: (1) les difficultés de consultation du FEW liées à ses spécificités typographiques et l'intérêt de développer des polices de caractères sur mesure, (2) les différentes stratégies d'encodage envisageables pour un dictionnaire qui emploie un système de transcription qui lui est propre, (3) l'évolution nécessaire de la mise en forme d'un contenu lexicographique lors de sa migration d'un format analogique à un format numérique.

MÜLLER, Caroline & SERAFIN, Stefan & REITER, Martin : El DEMel – Del fichero medieval a la publicación de acceso abierto en la web.

El objetivo del proyecto Diccionario del Español Medieval electrónico (DEMel), apoyado por la Fundación DFG (Sociedad Alemana de Investigación), consiste en facilitar el acceso, a través de internet, al archivo de datos del español medieval, resultado de más de 25 años del trabajo científico realizado en la elaboración del Diccionario del Español Medieval (DEM) en la Universidad de Heidelberg. Se trata de un archivo de datos extenso, con base científica, lematizado y estructurado semánticamente. Consiste de textos de más de 600 obras o colecciones de textos y documentos del siglo X hasta inicios del siglo XV. Contiene cerca de 865.000 fichas, de las cuales 650.000

corresponden a las fichas, en las que se encuentran anotadas las formas de palabras medievales con el contexto de su uso respectivo, la fuente del extracto ejemplar y su datación, así como p. ej. informaciones gramaticales, semánticas y etimológicas. La suma total de formas documentadas de un lema superarán el millón de ejemplos, ya que buena parte de estas fichas presenta no solo una sino más formas documentadas de un mismo lema. El leuario consta de cerca de 33.000 lemas. Medante la digitalización y el acceso electrónico de este archivo se asegurará el material a largo plazo y se creará un acceso libre a la comunidad científica internacional. La concepción del proyecto consiste en diferentes etapas, de las cuales algunas ya han sido finalizadas (p. ej. la digitalización del fichero, el desarrollo de una estructura de la base de datos etc.). 2 En la presente comunicación nos centraremos en las tareas actuales y las del futuro como p. ej. la interfaz de usuario en la página web para la obtención de las informaciones introducidas antes en la base de datos. Como introducción daremos muestras del material con el que trabajamos para mostrar su diversidad y para presentar luego el camino del material en papel hacia una obra digital. Queremos explicar también los problemas con los que se afronta un proyecto de este tipo y permitimos echar unas vistas a una posible evolución del DEMel en el futuro. Tratándose de un proyecto de lexicografía histórica unido a las nuevas tecnologías, el DEMel encaja perfectamente con la sección no 6 “Lexicographie”. Es único en su ámbito ya que hasta ahora no se ha publicado un archivo tan completo y elaborado del español medieval: representa el caudal léxico desde los inicios de la lengua hasta comienzos del siglo XV. El material completamente lematizado se basa en un abanico muy amplio de textos, tanto de carácter literario como de otras fuentes. De ello, resulta el gran potencial que ofrecerá el DEMel no solo a filólogos sino también a investigadores de otros campos.

PLAZA PICÓN, Francisca del Mar & GINÉS, Antonio Cano : Locuciones verbales de naturaleza mitológica en algunos diccionarios monolingües generales de la lengua española.

Atendiendo a las principales características de los fraseologismos como marco teórico, observamos el tratamiento que ciertas unidades fraseológicas (UFS), basadas en la mitología grecolatina, ofrecen en algunos diccionarios monolingües generales de la lengua española. Para la conformación del corpus tomamos como punto de partida el Diccionario de la lengua española (DEL), pero también basamos nuestro análisis en otros diccionarios: Diccionario de uso del español (DUE) de María Moliner, el Diccionario del español actual (DEA) de Manuel Seco y el diccionario Clave (Diccionario de uso del español actual). Los principales problemas con los que nos enfrentamos en nuestro análisis son la inexistencia de unos límites precisos en la tipología de las unidades fraseológicas, la falta de criterios claros acerca de qué unidades son susceptibles de ser introducidas en un diccionario y la ausencia de uniformidad en el tratamiento que reciben en los diccionarios. Asimismo, “el elemento que mejor ilustra la falta de uniformidad en el tratamiento de la fraseología es la diversidad terminológica que se puede observar en los diccionarios para referirse a las unidades que forman parte de ella” (Bargalló et al. 1997-1998: 52). En este sentido, según se desprende de las Advertencias para el uso del DEL, el concepto de locución se encuentra inmerso en la denominación genérica de formas complejas, en el DEA se habla de unidades léxicas compuestas o locuciones y en el DUE de expresiones pluriverbales, mientras que en el CLAVE hallamos la denominación de locución. Para el DEL son formas complejas aquellas series de palabras que, combinadas de una determinada manera, expresan conceptos no interpretables mediante la simple adición de los significados de sus componentes. En el DUE (2007) la “locución” se define expresión pluriverbal de forma fija que se

inserta en el habla como una pieza única, 2 constituida por una oración simple o compuesta o una parte de oración”. En el DEA por unidades léxicas compuestas o locuciones se entienden aquellas en las que una palabra perteneciente a una categoría establece combinaciones estables con otras palabras, lo cual le permite situarse en otras categorías distintas de la suya. Por ello será su funcionamiento el que marque su tipología (nominal, verbal, etc.). En el CLAVE la locución es una “combinación fija de palabras que forman un solo elemento oracional y cuyo significado no es siempre el de la suma de significados de sus miembros” (2002: 9). Como principio metodológico y con objeto de adaptarnos a los límites de tiempo propios de este trabajo acotamos las unidades fraseológicas relativas a la mitología grecolatina que en el DEL aparecen recogidas bien como locuciones verbales, para proceder posteriormente al análisis comparativo con los otros diccionarios ya citados. Este estudio cuyo detalle será objeto de la comunicación nos permite concluir que no existen criterios uniformes en los diccionarios consultados. Encontramos que en un mismo diccionario construcciones con similares características son tratadas de diversa manera: unas veces como locuciones y otras no. Además, como afirmaba García-Page (2004: 274), “No deja de ser arbitrario (...) el hecho de que una determinada locución aparezca registrada en unos diccionarios y no en otros, (...)”. En general, comprobamos que no existen ni se especifican criterios coherentes y exhaustivos a la hora de seleccionar y catalogar las UFS.

PUCHE LORENZO, Miguel Ángel : Vigencia, ampliación y pérdida léxica a través de las ordenanzas de Murcia del siglo XVII.

El estudio del léxico de la vida cotidiana de áreas periféricas del castellano, como es la correspondiente al antiguo solar del antiguo Reino de Murcia, nos ha proporcionado interesantes datos, tal como pudimos mostrar a partir del estudio de las ordenanzas medievales que se han conservado. La constatación de un amplio corpus léxico, fiel reflejo de los vaivenes políticos y socioculturales de ese territorio, nos ofrece la posibilidad de analizarlo diacrónicamente y comprobar cómo fue su evolución dentro de una sociedad cada vez más castellanizada. Dado que conocemos su comportamiento durante la Edad Media, podemos plantearnos ahora qué sucedió o cómo fue su devenir. Por ello, nuestro objeto de estudio será ahora las Ordenanzas del Campo y la Huerta, aprobadas por Carlos II e impresas en la ciudad Murcia en el año de 1695, sin olvidar la consulta de otras publicadas en el siglo XVI, pero de menor extensión. De hecho, los cambios sociales y económicos acaecidos en ese lapso temporal, convierten el código legislativo presentado en un rico testimonio del léxico referido a los usos sociales de una comunidad de hablantes concreta. El análisis del vocabulario que se recoge en las páginas de estas Ordenanzas nos permitirá comprobar la vigencia, en determinados ámbitos, de testimonios medievales que dotan de una peculiar caracterización lingüística a la interacción comunicativa de este territorio. Así mismo, la evolución de la sociedad trajo consigo la creación de nuevas voces que quedaron registradas en estos corpus legislativos que intentaban organizar la vida y las actividades en las ciudades. Pero también hubo otras que, en justa medida, desaparecieron y su estudio nos ayuda a conocer los usos sociales del pasado. Dado que son escasos los trabajos que han mostrado interés por el léxico de estas áreas periféricas, los resultados que se obtengan nos darán a conocer interesantes datos que nos permitirán establecer con mayor precisión esa parcela de la historia de nuestra lengua, en general, con toda su complejidad y otra marcada dialectalmente, en particular. Muestra de todo lo expuesto, será lo que se registrará en el campo de los arabismos, pues muchos seguirán vigentes, desde la Edad Media, durante ese periodo,

incluso hasta la actualidad, por haberse convertido en claro ejemplo de un campo de especialidad. Aunque esa área especializada solo esté registrada allí o en zonas con similitudes geográficas u orográficas, tal como sucede en el caso del campo semántico del “agua”. En otros casos, sin embargo, desaparecerán porque o bien lo ha hecho el referente que designaban o bien han sido sustituidas por una voz de origen latino. Un caso similar se observará en los orientalismos, tan frecuentes en esta zona de tránsito. Junto a estos fenómenos que nos sugieren una intensa investigación acerca de los vaivenes que se experimentan en la historia del léxico español, se hallarán neologismos que se difunden en virtud de la nueva organización social y que, con el paso del tiempo, serán valorados como dialectalismos. El interés y el valor de este tipo de trabajos es amplio dado que se podrán sentar las bases para la elaboración de una obra lexicográfica, de carácter histórico, relativa a la zona geográfica presentada. Debemos indicar, en primer lugar, que no existe ningún corpus digitalizado, hasta la fecha, sobre el que realizar este estudio; no se ha redactado ningún diccionario de tipo histórico de ámbito murciano; no hay, salvo en algún caso aislado y lejano en el tiempo, obras lexicográficas realizadas con criterios sólidos y científicos; los casos recogidos en las sucesivas ediciones del DRAE, ahora DLE, presentan vaivenes en el tiempo y, en ocasiones, no resulta apropiada la marcación diatópica que las acompaña. En consecuencia, el estudio que presentamos parte de documentación original, hasta ahora no tratada en el seno los estudios lingüísticos, y para el tratamiento léxico solo disponemos de los recursos, escasos, disponibles a través de las obras académicas, principalmente, ya sean en línea o en papel, así como de otras obras lexicográficas no académicas.

RENDERS, Pascale : Lexicographie historique et Web sémantique : intégration de la dimension étymologique dans le modèle OntoLex-Lemon.

La lexicographie se trouve actuellement en pleine transition vers le numérique. À côté de la rédaction informatique de dictionnaires en ligne et de la numérisation d'ouvrages imprimés, des initiatives européennes telles que le projet LIDER (<http://lider-project.eu>) ou l'action COST European Network of e-Lexicography (ENeL, <http://www.elexicography.eu>) ont conduit à l'élaboration de modèles informatiques qui permettent aujourd'hui aux ressources lexicographiques d'intégrer, partiellement ou totalement, le Web sémantique et, en particulier, le réseau constitué par le Linguistic Linked Open Data (<http://linguistic-lod.org/lod-cloud>). Très récemment, le projet Elexis (European Lexicographic Infrastructure), qui a débuté en février 2018, s'est donné pour objectif de favoriser la coopération entre les équipes de recherche afin notamment de développer des standards, des méthodes et des outils permettant la mise en réseau des ressources lexicographiques et l'accès libre à leurs données (<http://www.elex.is>). Parmi les standards qui sont proposés à la communauté des lexicographes se trouve le modèle OntoLex-Lemon (<http://www.w3.org/2016/05/ontolex/>), qui a déjà été testé sur plusieurs dictionnaires en cours d'informatisation (voir par exemple Declerck/Wandl-Vogt 2015). Ce standard n'avait pas pour objectif premier de permettre la mise en réseau de dictionnaires, mais d'enrichir les ontologies du Web sémantique avec des informations d'ordre lexical. Cet objectif explique pourquoi le modèle ne permettait pas initialement (et ne permet pas facilement aujourd'hui encore) de représenter toute la complexité des informations présentes dans les ressources lexicographiques. Son application en lexicographie a rapidement donné lieu à des améliorations permettant d'y ajouter des informations d'ordre par exemple morphologique ou syntaxique. En revanche, la dimension étymologique et plus globalement historique est encore loin d'être prise en

compte dans le modèle, même si quelques propositions ont été émises pour y ajouter la possibilité de renseigner des étymons lorsque les dictionnaires contenaient cette information (cf. Declerck/Wandl-Vogt 2015). La dimension historique reste, de façon générale, peu présente au sein du Linked Open Data (voir par exemple Tittel/Chiaros 2018, qui tente d'intégrer le DEAF dans ce réseau). La communication que nous proposons au XXIXe Congrès international de linguistique et de philologie romanes poursuit deux objectifs. Le premier est d'étudier la possibilité d'intégrer les ressources de référence des langues romanes dans ce réseau, en distinguant les ressources proprement lexicographiques de celles qui s'apparentent davantage à des thesauri (le FEW et le LEI, par exemple) et de celles qui ne sont pas lexicographiques à proprement parler (les atlas linguistiques, les glossaires ou encore les corpus annotés). Le second objectif est de proposer une amélioration du modèle OntoLex en y ajoutant la dimension « étymologie-histoire du mot » telle qu'elle est pratiquée en linguistique romane. Les deux objectifs sont étroitement liés car c'est l'intégration dans le réseau d'ouvrages tels que le FEW qui permet l'ajout de cette dimension étymologique, tandis que l'ajout de cette dimension permet en retour la mise en réseau d'un ensemble d'autres ressources romanes, qu'elles soient lexicographiques ou non. L'exposé présentera la méthodologie que nous proposons pour atteindre ces deux objectifs. Après une synthèse des principes qui guident le Web sémantique, nous expliquerons le modèle OntoLex et la façon dont nous pensons l'améliorer de façon à y intégrer pleinement les dimensions historique et étymologique. Il s'agit d'une part d'ajouter au modèle des propriétés permettant de renseigner une filiation entre les unités lexicales et, d'autre part, de créer des ontologies susceptibles de servir de ressources externes auxquelles chaque dictionnaire pourra se reporter. En nous focalisant sur les ressources galloromanes, nous expliquerons comment le FEW peut, notamment, fournir un excellent point de départ pour la création d'une ontologie à visée étymologique disponible pour toute ressource du Web sémantique. L'amélioration du modèle OntoLex et la mise à disposition de nouvelles ontologies doit ensuite permettre de mettre en réseau les ressources lexicographiques romanes. Nous verrons en pratique comment divers dictionnaires et atlas des langues romanes pourront, au moyen de ces nouveaux outils, intégrer le Web sémantique tout en l'enrichissant d'informations d'ordre historique et étymologique.

ROBECCHI, Marco : La collaboration mutuelle entre DocLing et FEW.

Notre proposition s'insère dans le projet La « scriptologie » lexicale : les cas du domaine oïlique oriental financé par le FNS et dirigé par H. Carles et M. Glessgen. Ce projet voit la collaboration entre le Centre de Dialectologie et d'Étude du Français Régional de l'Université de Neuchâtel, l'Université de Zurich et le laboratoire ATILF de Nancy, notamment la rédaction du *Französisches Etymologisches Wörterbuch*. Le but du projet est l'étude de la variation diatopique du lexique employé dans les textes surtout documentaires d'une région donnée, le Nord-Est de la France (Champagne, Lorraine, Bourgogne et Franche-Comté), tout au long du Moyen Âge. Le choix de cette aire géographique est essentiellement lié à la disponibilité du corpus de textes documentaires rédigés aux 13e et 14e siècles qui constitue la base de données DocLing. L'interrogation informatique de ce corpus, qui sera suivie par l'analyse de quelques œuvres non documentaires afférentes à la même région, nous a permis de traiter – pour l'instant – quelques 400 régionalismes. L'étude du régionalisme, « un lexème dont la forme et/ou le sens se caractérise par une diffusion régionale à l'intérieur de l'espace de la langue en question » (Glessgen 2016, 3) revêt un intérêt particulier pour l'étude, plus ample, de l'histoire de la langue française et de sa constitution à partir des premier

témoignages écrits. Ce type d'analyse nous a déjà permis d'obtenir des données supplémentaires de première importance : à côté du marquage diatopique des mots considérés, nous avons pu recueillir un bon nombre d'informations relatives à leur marquage diachronique et diastratique. Ce marquage ne peut être considéré qu'au sein d'une analyse différentielle qui tienne compte du rapport des documents régionaux avec le reste de la production écrite oïlique. Les outils nécessaires pour ces interrogations sont les dictionnaires d'ancien et de moyen français, le Godefroy, le Tobler-Lommatzsch, le DEAF, le DMF et le récent dictionnaire de Matsumura ; on y ajoute les dictionnaires de latin médiéval de Du Cange et de Niemeyer et les bases de données de l'ARTEM et des Chartae Galliae. Même si, de manière générale, ils ne donnent pas d'informations explicites sur le degré de régionalité d'un lexème, l'analyse comparative des attestations enregistrées peut donner une idée de l'extension diatopique (et diachronique) d'un mot donné. Dans le cadre de cette enquête, l'utilisation du FEW devient incontournable. La masse de matériaux employés pour la rédaction de ce dictionnaire donne un aperçu global de la langue française dans sa variation diatopique et diachronique. Cependant, à l'époque de rédaction du FEW une analyse si précise de l'écrit documentaire n'avait pas encore été effectuée : par conséquent, bon nombre de données ont échappé aux rédacteurs et sont absentes du dictionnaire. Bien entendu, le FEW permet presque toujours de clarifier les problèmes étymologiques sous-jacents dans l'identification souvent épineuse des mots attestés par les documents et qui ont été traités de manière imprécise par les autres dictionnaires. Par exemple, le rasel "outil pour mesurer les grains" (< *RASARE), est à distinguer du resal "unité de mesure des grains" (< RASUS), même si l'aboutissement formel et le contexte provoquent souvent une confusion entre les deux formes ; le poil "couche d'herbe" (lorr., bourg.), absent des dictionnaires sauf Gdf "herbe", est à rattacher à PILUS (FEW 8, 510b), qui enregistre des att. des patois de Saint-Pol et de Saône-et-Loire qui éclaircissent la nature de ce mot, autrement peu évident. Dans cet échange réciproque, les DocLing peuvent compléter la connaissance du français fournie par le FEW, préciser l'extension régionale et l'extension dans le temps, indiquer les usages spécifiques. Le vb. festüer (wall./pic., frcomt.) "renoncer à ses biens" (< FESTUCA, FEW 3, 485a) manque au FEW, qui n'enregistre que le vb. préfixé effestuer (wall. 13e s.). Le mot debit s.m. (lorr., frcomt.) "sorte de transaction liée à une question juridique" (< DEBITUM, FEW 3, 22a), bien attesté par nos documents bien que dans un emploi diastratiquement délimité, n'est pas enregistré par FEW, qui ne considère que le latinisme debite s.f., hapax du 15e s. Enfin, l'adj. païable (< PACARE, FEW 7, 457b) est attesté dep. 1250 par DocLing (avec une première attestation latine en 1232), alors que FEW mentionne la première attestation en 1255 (d'après Godefroy). Le FEW est une œuvre accomplie ; les lettres A et B ont fait (et font) l'objet d'une refonte. On expérimente actuellement la possibilité de l'informatiser pour le rendre disponible à une « lecture transversale » (Renders 2015, 7) afin qu'il puisse être interrogé en combinant certains paramètres spécifiques (espace, temps, niveau, etc.). En outre, la version informatisée, sans toucher à la structure et à la version finale voulue par von Wartburg, permettra d'intégrer les données supplémentaires et les nouveaux matériaux déjà publiés. On peut finalement comprendre, dans cette perspective, l'apport fondamental de l'étude lexicale des DocLing pour un outil indispensable comme le FEW, sans négliger la contribution que ce dernier peut toujours donner à toute nouvelle enquête linguistique sur la langue française.

SÁNCHEZ LANCIS, Carlos : Lexicografía y pragmática en español y catalán: Tabú y eufemismo en el DLE (2014) en contraste con el DIEC2 (2007).

A partir del análisis de la presencia y la evolución en los diccionarios y en los corpus académicos de algunos pares de vocablos y expresiones en Sánchez Lancis (2016), se constató que el diccionario es un claro reflejo de una realidad social a través del tabú y el eufemismo. En ciertas ocasiones, un vocablo adquiere unas connotaciones semánticas marcadas no deseadas (tabú) y es sustituido por otro más neutro o no marcado (eufemismo) para designar, en teoría, el mismo referente. El DLE (2014: s.v. eufemismo y tabú) define eufemismo como la 'manifestación suave o decorosa de ideas cuya recta y franca expresión sería dura o malsonante', mientras que tabú corresponde a la 'condición de las personas, instituciones y cosas a las que no es lícito censurar o mencionar' (el DIEC2 (2007: s.v. eufemisme, tabú y mot1) recoge también definiciones muy parecidas, pero destaca especialmente la de mot tabú como 'Mot que es considera impropri d'emprar en públic o d'escriure per raons religioses, de superstició, de gust social, etc.'). Ciertamente, para que exista uno o más eufemismos estos tienen que relacionarse con la existencia de un vocablo considerado por los hablantes como tabú. Se trata, en definitiva, de cambios semánticos que presentan variación tanto diacrónica como diatópica, por lo que los corpus permiten complementar y contrastar la información aportada por los diccionarios en el estudio de este tipo de evolución; así como los mismos diccionarios se convierten también en corpus indirectos, ya que reflejan no solo la existencia sino también la percepción social de esta clase de elementos por parte de los lexicógrafos que los compendian. Este tipo de cambio semántico, que se produce de forma bastante frecuente en la lengua, es reflejado, con mayor o menor acierto, por parte de los diccionarios, en general, pero sobre todo por los diccionarios de carácter más académico, en particular. Los objetivos del presente estudio son varios. En primer lugar, se va a completar el análisis, a partir del trabajo de Sánchez Lancis (en prensa), de las diferencias existentes entre la edición del Diccionario de la lengua española (DRAE: 2001) y la nueva edición más reciente (DLE: 2014) de la RAE y la ASALE , para constatar en qué forma (tipo de marcas lexicográficas, comentarios específicos, definiciones, etc.) estas obras reflejan el eufemismo (y el correspondiente tabú), y en qué medida se ha producido algún cambio en las entradas del nuevo diccionario académico debido a un cambio en la percepción social de estos vocablos por parte de los hablantes. Nos interesan tanto los eufemismos que se utilizan para evitar términos que hacen referencia a realidades desagradables (acciones fisiológicas, etc.), como muy especialmente los que se emplean para manipular la realidad (el relato de la guerra, de la crisis económica, etc.), ya que estos son los que han proliferado en mayor medida frente a los primeros en los últimos años. Por ejemplo, el DRAE (2001) distingue 72 eufemismos frente a 91 palabras tabú mediante las marcas "eufem." y "malson." respectivamente, mientras que el DLE (2014), con el mismo procedimiento, destaca 88 eufemismos frente a 270 palabras tabú, con un incremento muy notable de estas últimas formas, que merece la pena ser estudiado. En segundo lugar, se va a realizar el mismo tipo de análisis con el Diccionari de la llengua catalana (DIEC2: 2007), sin dejar de lado como referentes el Diccionari CatalàValencià-Balear (DCVB) y el más reciente Diccionari descriptiu de la llengua catalana (DDLC). Por ejemplo, el DIEC2 (2007) recoge solamente 15 lemas con el vocablo "eufemisme" en su definición, frente a 42 posibles palabras tabú identificadas mediante la marca "VULG." ('utilizado vulgarmente'), aunque probablemente habría que tener también en cuenta bastantes de los vocablos caracterizados con la marca "POP." ('utilizado popularmente'). Por último, se contrastan los diccionarios de ambas lenguas, español y catalán, para ver hasta qué punto las palabras que corresponden a un mismo campo semántico son consideradas de forma idéntica o diferente en relación con los conceptos de eufemismo

y tabú. Como complemento, se realiza también un estudio de la evolución del significado de alguno de los vocablos analizados en los diferentes bancos de datos académicos (CORPES XXI, CREA y CORDE, y también CTILC), así como de su reflejo en los diccionarios que se recogen en el NTLLE.

SÁNCHEZ-PALOMINO, María-Dolores : As linguas románicas e a romanística desde a perspectiva dos lexicógrafos galegos.

O amplo uso do dicionario na sociedade, e moi en particular no ámbito educativo, fai desta obra un instrumento especialmente relevante á hora de vehicular determinadas ideas ou concepcións. A este valor engádeselle, no caso das obras académicas, a súa transcendencia nos procesos de estandarización lingüística. Por último, cómpre salientar a importancia crecente da disciplina lexicográfica dentro dos estudos lingüísticos desde hai algunhas décadas. Con estas tres dimensións se relaciona o traballo que presentamos. A través dunha análise da tradición lexicográfica galega, en concreto da tradición dicionarística (van alá aproximadamente cento cincuenta anos), poderemos observar a evolución da visión que os lexicógrafos, ao principio afeccionados, hoxe plenamente formados na disciplina lexicográfica, foron tendo do feito románico, o que equivale a dicir a visión que foron transmitindo á sociedade, ás aulas etc. A análise, que combina a visión das linguas románicas no seu conxunto co estudo da situación na lexicografía dunha lingua en particular como é o galego, enmárcase, ademais, no ámbito da historia desa disciplina e das linguas románicas en xeral, de xeito que ten tamén un interese propiamente lingüístico e historiográfico. O corpus de partida estará, pois, constituído polos dicionarios monolingües de galego, as enciclopedias publicadas nesta lingua (Gran Enciclopedia Galega, Enciclopedia Galega Universal), así como tamén os dicionarios correspondentes a practicamente toda a etapa prenormativa (ata os anos 80 do século pasado) nos que a entrada se presentaba habitualmente en lingua galega pero as equivalencias ou definicións se redactaban en castelán. Serán tidas en consideración tanto as obras impresas como aquelas que se ofrecen en liña. Canto aos elementos que se examinarán, terase en conta o tratamento das diversas linguas e variedades románicas, así como de todos aqueles fenómenos polos que se ven afectadas e nos que pode existir unha mención a elas (gheada, seseo, lenición etc.), sen esquecer o tratamento que recibe a propia romanística. Abordarase a inclusión ou non como lemas das voces correspondentes, a información gramatical que se achega sobre elas, as definicións, a marcación diatécnica, os sinónimos e remisións, a eventual fraseoloxía, as observacións e a exemplificación, na que se poden atopar interesantes datos e mencións. En resumo, serán obxecto de análise aqueles elementos que conforman o artigo lexicográfico e que poden proporcionar información sobre a consideración do mundo románico e da romanística.

SITA FARIAS, Virginia : Diccionario, estándares nacionales y autoafirmación lingüística: Desafíos y perspectivas de la lexicografía hispanoamericana actual. Los casos del *Diccionario del español de México* – DEM (2010) y el *Diccionario integral del español de la Argentina* – DIEA (2008).

A propósito de la relación que Weinrich (1985) establece entre el diccionario y su época, Bugueño Miranda (2016: 89-90) subraya que entre las preocupaciones más inmediatas de la lexicografía –y más específicamente la de lengua española– están el diseño de diccionarios de aprendices [learner's dictionaries] y la descripción del potencial combinatorio de las palabras. En lo que a la lexicografía

hispanoamericana concierne –y atendiendo a las consideraciones de Zimmermann (2010: 54 ss.) sobre la hispanofonía y la aceptación de una cultura lingüística pluricéntrica–, habría que añadir a la lista de tareas pendientes la elaboración de diccionarios integrales nacionales. Hasta esta fecha se han concluido apenas los proyectos de México (DEM 2010) y de Argentina (DIEA 2008). En vista de la ineludible fragmentación terciaria del español americano (cf. Lüdtke 2007) y el subsecuente estado de “orfandad normativa” de la comunidad hispanoamericana (cf. Farias 2018a; 2018b), los proyectos de diccionarios integrales responden, pues, por una parte, a la necesidad apremiante de descripción de los estándares nacionales/regionales en América. Por otra parte, y como consecuencia de lo anterior, los diccionarios integrales, al contribuir en la fijación de los estándares de cada país, brindarían asimismo un aporte significativo a lo que Oesterreicher (2002) denominó “autoafirmación lingüística nacional” de las comunidades hispanoamericanas. La lexicografía integral abarca, por lo tanto, una perspectiva lingüístico-normativa y otra políticoideológica, que aunque pueden deslindarse, han de interpretarse complementariamente. De este modo y sobre la base de las consideraciones anteriores, el objetivo que aquí se plantea es analizar los dos diccionarios integrales ya publicados –DEM y DIEA– tomando en cuenta tres dimensiones: 1. Dimensión descriptiva: Oesterreicher (2002: 276 ss.) hace hincapié en que el espacio variacional de una lengua histórica se constituye a partir de la delimitación del estándar (cf. tb. Coseriu 1982), de manera que éste funciona como una suerte de “fuerza ordenadora”. En el ámbito (meta)lexicográfico, y siendo el fin principal del diccionario la descripción lingüística, uno de los principales escollos consiste justamente en determinar la variedad –o lengua funcional, en términos coserianos, considerada como estándar– que deberá servir de punto de referencia para el registro de las formas pertenecientes a otras variedades diasistémicamente diferenciadas (cf. Kilian 2005: 116 ss.). Ahora bien, la descripción lingüística en el ámbito de la lexicografía integral hispanoamericana se enfrenta a un doble desafío –y por ende adquiere también una doble importancia–. Por un lado, los diccionarios integrales cargan con la responsabilidad de describir una “variedad nacional”, para emplear las palabras de Lara (2015: 29-31). Ello supone, claro está, la delimitación –y consecuentemente la fijación (cf. ad infra)– de una “ejemplaridad nacional” –o estándar– a partir del cual se hace posible determinar lo “diferencial” en el ámbito de una comunidad lingüística. De esa manera, los diccionarios integrales no sólo contribuirían a describir adecuadamente la variedad nacional en cuestión, sino también a distinguir, según la propuesta de Oesterreicher (2002: 281 ss.), entre estándar regional y variación diatópica. Por otro lado, los diccionarios integrales podrían representar asimismo un valioso aporte en la descripción del pluricentrismo del español. En este sentido, Rivarola (2006: 106) sostiene que a una lengua pluricéntrica debería corresponder una descripción de carácter pluricéntrico. Ello, sin embargo, sólo es posible si se conocen (y se han descrito) todas y cada una de las variedades estándares de la lengua en cuestión –lo que para el español sigue siendo una desiderata (cf., p. ej., Greußlich 2015)–. 2. Dimensión normativa: La normatividad, por una parte, corresponde a un anhelo genuino de la comunidad lingüística en el nivel del saber idiomático y, por otra, representa un aspecto inherente a la naturaleza del diccionario –ya sea implícita o explícitamente– (cf. Farias 2018a: 163- 164). En lo que atañe al problema de la norma en el ámbito hispánico, Oesterreicher (2002) y Rivarola (2006), por ejemplo, coinciden en que una codificación exclusivamente monocéntrica ya no se correspondería con la realidad pluricéntrica del español. Sin embargo, y como forma de paliar la ya mencionada ausencia de descripciones sistemáticas de los estándares americanos, Pöll (2012) sigue defendiendo el ideal de una norma panhispánica compatible con el carácter pluricéntrico del español. Los diccionarios integrales, en ese sentido, cumplirían dos funciones. Desde un punto de vista estrictamente lingüístico (cf. ad supra), a la vez que describen una “variedad nacional” contribuyen

asimismo a la fijación de la norma ejemplar o estándar nacional/regional. Ya desde una perspectiva político-ideológica (cf. ad infra), al fijar los estándares nacionales/regionales, los diccionarios integrales podrían llegar a convertirse en herramientas de orientación normativa, ofreciéndole a la comunidad lingüística una norma con la que ésta se identifica –terminando por restarle fuerza a la llamada “norma panhispánica”–. En razón de lo expuesto, parece prudente sopesar los argumentos que Fajardo (2011: 60 ss.) evoca en contra del valor normativo del DEM y el DIEA. El autor, sin lugar a duda, lleva razón al cuestionar la pertinencia de la descripción lingüística en esas obras. Para poner un ejemplo: el conjunto de textos seleccionados así como la metodología anticuada aplicada en la elaboración del corpus, fraguado en los años 1970, ponen en tela de juicio el valor descriptivo del DEM (cf., p. ej., Fajardo 2011: 61 ss.). De modo similar, Lauria (2015) y Adelstein (2016) admiten restricciones en el corpus original y señalan la necesidad de recurrir a fuentes complementarias en la elaboración del DIEA –lo que ha redundado en inclusión de léxico perteneciente a variedades ajenas a la(s) argentina(s) y en ausencia de marcas diatópicas relativas a las variedades regionales del país, comprometiendo la imagen de la lengua que el diccionario proyecta y haciendo, hasta cierto punto, inviable la delimitación del estándar frente a las demás variedades–. En ambos casos se trata, no obstante, de problemas metodológicos – o de consecución del fin principal de la obra– que evidentemente deben corregirse. En un principio, sin embargo, y al contrario de lo que defiende Fajardo (2011: 63 ss.), nada obsta para que los diccionarios integrales, en el plano (meta)lexicográfico, puedan –y deban– concebirse con el propósito de “convertirse en referencia normativa para su país”. 3. Dimensión político-ideológica: Bergenholtz y Gouws (2006: 2) remarcan que toda y cualquier decisión lexicográfica constituye al fin y al cabo una decisión política. Lo político, por lo tanto, rebasa los límites de la mera planificación lingüística [language planning], adquiriendo una “dimensión ideológica”, por así decirlo, que, naturalmente, abarca la política lingüística stricto sensu (intrínsecamente ligada a su dimensión normativa, cf. ad supra), pero es más amplia, ya que el contexto histórico, cultural y social de la publicación de una obra lexicográfica cualquiera determina, como se sabe, su “modo de ser” en cada época. Así pues, el diccionario integral –aunque sólo sea idealmente– al presentarle al usuario un instrumento de prescripción lingüística – ésta, a su vez, fundamentada en la descripción rigurosa de la(s) variedad(es) en cuestión a partir de sólida teoría lingüística y metalexicográfica–, podría responder a la demanda de “su época”, es decir, la “elaboración de normas por países [...] y su aceptación como tales en cada país” (Zimmermann 2010: 54) en favor de una “autoafirmación lingüística nacional” (cf. Oesterreicher 2002).

TITTEL, Sabine : La lexicographie historique et le Web Sémantique : La transformation du Dictionnaire étymologique de l'ancien français en Linked Data.

Depuis environ deux décennies, les grands dictionnaires scientifiques s'investissent pour la transformation de leurs produits en publications électroniques accompagnés de fonctions multiples. En même temps, de nouveaux produits dictionnaires se constituent comme projets numériques dès l'origine. Le paradigme établi pour la création, l'entretien et la publication des ressources lexicographiques numériques est l'emploi de XML comme format de données (et des formats apparentés, p.ex. XHTML). Toutefois, les publications des dictionnaires présentent ce que Tim Berners-Lee a nommé un data silo (cf. McCrae et al. 2012, p. 25). Le terme se réfère à une ressource linguistique dont l'accès est public, mais seulement dans un format propriétaire ou à accès limité. Les données d'un data silo ne peuvent être utilisées que dans la manière initialement prévue par leurs

créateurs et dans la manière facilitée par le format XML. Pour la plupart des ressources, cela se limite à une recherche par mot-clé et une recherche intégrale sur plusieurs ressources est souvent impossible. Une recherche dictée par une approche onomasiologique partant du sémantisme des mots est rendue difficile par cette configuration. Par exemple, trouver les mots français, occitans ou italiens du Moyen Âge désignant des états pathologiques ou des concepts de guérison associés à la théorie des humeurs est une tâche laborieuse. Avec les fonctions de recherche typiquement établies, la réussite de cette tâche présuppose la connaissance préalable du résultat-même de la recherche. Une solution au problème d'accès limité aux données d'un data silo – pas uniquement pour l'approche onomasiologique de la recherche – est la transformation des données en Linked Data (cf. Heath et Berners-Lee 2009) en utilisant les formats et les technologies du Web Sémantique. Cette transformation est possible, p.ex., par l'application du modèle OntoLex-Lemon (LExicon Model for Ontologies) pour la conversion des ressources lexicographiques. Toutefois, le processus de la transformation a avant tout été appliqué à des dictionnaires des langues modernes. Dans la communication présente, nous souhaitons présenter l'application de ce modèle à une ressource linguistique historique, dans le cadre de la transformation des données du Dictionnaire étymologique de l'ancien français – DEAF (Baldinger et al. 1971–). Nous commençons par une définition de Linked Data et une présentation des formats et vocabulaires impliqués. Ce sont notamment le modèle de graphe RDF (Resource Description Framework) pour la description des ressources et de leurs métadonnées et ce sont les ontologies et vocabulaires utilisés pour la modélisation des données du DEAF dont le vocabulaire OntoLex-Lemon. Puis, nous montrons la modélisation d'un article type du DEAF et nous présentons la conversion automatique des données XML du DEAF en RDF. Finalement, nous discutons les difficultés de modélisation dues à l'état médiéval de la langue et à la structure complexe des articles du DEAF (cp. Tittel 2018, p. 352-366). Depuis environ dix ans, le terme Linked Data se réfère à un paradigme pour la publication des jeux de données dans le Web Sémantique ; il désigne un groupe de formats et de technologies standards développés pour la création et la publication des données structurées, de façon à permettre leur traitement automatique (cf. Bizer et al. 2009). Les données sont annotées avec des informations sémantiques – dans le sens informatique –, de manière à établir des liens avec d'autres données. En informatique, et en particulier dans le domaine du Web Sémantique, la sémantique est l'expression explicite des relations entre les données ajoutant aux données des informations sur leur contenu (cf. Hitzler et al. 2008, p. 13). La liaison de la chaîne avec les informations et est une annotation sémantique. Pour différencier la sémantique informatique de la sémantique comme branche de la linguistique nous nous référons, par la suite, à la dernière comme lexico-sémantique. Il y a quatre principes à l'établissement de Linked Data : (1) utiliser des URI (Uniform Resource Identifier) pour chaque information, (2) utiliser HTTP (Hypertext Transfer Protocol) pour les URI, (3) utiliser les standards W3C (World Wide Web Consortium) comme RDF, et (4) inclure des liens à d'autres ressources. Suivre le paradigme de Linked Data présente les avantages suivants : l'interopérabilité structurelle (formats et langage de requête communs), l'interopérabilité conceptuelle (vocabulaire commun), l'accessibilité (accès uniforme, données utilisables à l'aide de protocoles Web standard sans logiciel supplémentaire, etc.) et l'intégration des ressources (accès sémantique à travers plusieurs ressources). Le vocabulaire OntoLex-Lemon a été développé pour créer une interface entre les ressources lexicographiques et les ontologies. Le vocabulaire fournit des moyens pour la modélisation des informations d'une entrée typique de dictionnaire : d'une part, les métadonnées, c'est-à-dire le créateur de l'article, le droit d'auteur, la date de la création, etc., et d'autre part, les informations linguistiques ; ce sont l'entrée (le lemme), la partie de discours, les formes déclinées, les variantes graphiques (pour les états

historiques des langues), les relations avec des expressions multi-mots, et les sens et concepts des mots. En utilisant OntoLex-Lemon (en combinaison avec d'autres ontologies), nous montrons la conversion de l'article FESTRE f. du DEAF. Cet article est d'une complexité et d'une longueur moyenne, afr. festre ayant onze variantes graphiques et désignant un conduit, une sorte de flûte, une fistule et d'autres choses. En outre, festre figure dans les collocations gote festre, une maladie, et cassia fistre, le fruit du cassier. Nous avons converti l'article type d'une manière manuelle afin d'adapter le modèle aux besoins du DEAF et pour définir le format cible. Pour la conversion automatique des données, nous avons ensuite développé six scripts XSLT. Le but de ces scripts est de pouvoir convertir environ neuf mille articles en RDF d'une manière automatisée. Lors du processus de modélisation, plusieurs problèmes sont apparus, dont celui de l'annotation lexicosémantique – un des premiers enjeux de la modélisation : Pour cette annotation, il a fallu établir des liens entre les sens des mots, p.ex. “fistule” ou “fruit du cassier”, et les entités d'une ontologie représentant (une partie de) tout ce qui existe dans le monde. La base pour l'établissement de ces liens est constituée par les concepts de la réalité extralinguistique représentés par les mots. DBpedia, la version Linked Data de Wikipedia, peut servir comme modèle (partiel) du monde contemporain. Cependant, la grande différence entre le monde médiéval et le monde contemporain constitue un obstacle à l'établissement de liens. Le second enjeu est l'identification des scriptae de l'ancien français, p.ex. le picard et l'anglonormand. Nous utilisons la norme internationale ISO 639 qui définit des codes pour représenter des noms de langues, p.ex. ‘fr’ pour ‘français’ et ‘fro’ pour ‘ancien français’. Toutefois, des codes pour les scriptae de l'ancien français n'existent pas. Il en résulte une perte d'information regrettable : La caractérisation d'une variante graphique comme, p.ex., picarde, une information précieuse contenue dans les données source XML, manque forcément dans la version RDF. Le but de notre communication est de montrer une nouvelle approche des riches ressources lexicographiques : La transformation des données en Linked Data permet une réutilisation qui dépasse les limites des data silos, c'est-à-dire les limites des publications électroniques conventionnelles. Nous sommes convaincues que les efforts investis se justifient par rapport à la valeur ajoutée aux jeux de données qui sont les résultats des travaux lexicographiques précieux.

VEGA Y VEGA, Jorge Juan : L'article du verbe *être* dans le *Trésor de la Langue Française*.

Les études linguistiques sur le verbe être de la langue française restent encore assez fragmentaires. Un effort en vue d'en proposer une approche unitaire serait vivement salubre. En ce sens, il nous semble nécessaire de considérer la nature et le type d'informations que les dictionnaires ont proposés pour la définition du verbe. Notre approche ne serait pas tellement celle du lexicographe expert dans la présentation et portée des articles du dictionnaire, mais plutôt celle de l'observateuranalyste qui vise à concevoir un ensemble structuré et organique des connaissances concernant le verbe être, à partir cette fois-ci des informations que ces mêmes articles ont thésaurisées. L'analyse du détail portera tout particulièrement sur le cas du Trésor de la Langue Française informatisé (TLFi). Nous y retrouvons en effet l'article de dictionnaire le plus complet que nous ayons pu étudier. Produit d'une longue réflexion qui a su tenir compte de la tradition dictionnaire, cet article, très original quant à son architecture textuelle, est sans doute une inestimable et très dense référence lexicographique pour mesurer l'étendue et la richesse du vocabulaire. Toutefois, comme le dit O. Soutet : « à s'en tenir au seul verbe être, la polysémie n'est pas traitée par la médiation des opérateurs métalexigraphiques essentiels, ‘extension de sens’, ‘restriction de sens’, ‘métonymie’ et ‘métaphore’, mais selon un

schéma mixte associant une approche sémantique, mais conceptuellement faible (avec une opposition fondamentale entre un sens ‘fort’ et des sens perçus comme moins forts) et une approche syntaxique, mais assez peu élaborée (avec distinction d’emplois intransitifs et d’emplois dans des expressions dites ‘binaires’ [...]) Qu’on nous comprenne bien : nos remarques ne visent nullement à condamner l’article du TLFi, clair et accessible ; elles se bornent à souligner que son organisation s’appuie sur des notions le plus souvent préthéoriques (mais intuitivement accessibles) » (2016, 33-34). Pour en obtenir, donc, une vision d’ensemble, nous essayerons de comprendre et d’expliquer comment l’article du verbe être a été présenté et structuré dans le TLFi, et quels ont été les critères généraux et les subdivisions particulières qui ont guidé le choix des lexicographes. Nous y signalerons au passage certaines caractéristiques et quelques singularités, associées aux diverses descriptions du lemme : acceptions, distributions, renvois, lacunes, etc. Notre souci majeur est de soumettre toutes ces données à un profond questionnement sémantique : quelles sont les diverses significations du verbe être ? Et comment peuvent-elles se combiner entre elles sur un plan lexicographique ? Sur ce point, les éclairages contrastifs venant d’autres langues romanes, telles que l’espagnol (ser, estar, haber, etc.) ou l’italien (essere, stare, esserci, starci, etc.), et même celui de l’ancien français (estre vs ester), ainsi qu’un renvoi à la langue mère, nous sembleront particulièrement révélateurs. Notre objectif final est de pouvoir dresser (selon un classement décimal) un tableau général des connaissances lexicographiques réunies jusqu’ici sur le verbe le plus polyvalent et utilisé de la langue française. À partir donc de cette si riche et dense variété de données, nous essayerons de proposer – en y introduisant une composante cognitive nullement négligeable désormais – quelques réflexions sur un autre possible tableau distributif de la matière lexicale du verbe être.

VIARO, Mário Eduardo : A IMPORTÂNCIA DOS DICIONÁRIOS BILÍNGUES DE JERÔNIMO CARDOSO, PARA A CONFECÇÃO DE UM DICIONÁRIO ETIMOLÓGICO DA LÍNGUA PORTUGUESA.

Dentre todas as línguas românicas oficiais, o português é o idioma que tem mais indagações acerca de sua história. À busca de um dicionário etimológico que responda a questões sobre as primeiras ocorrências dos componentes de seu léxico, normalmente recorre-se a Cunha (1982, 1989, 2003, 2006), a Machado (1952-1977, 1981), a Houaiss & Villar (2001), hoje na forma *online* do Grande Dicionário Houaiss, e a Machado Filho (2013), quando não a Corominas. (1953) e a Corriente (2008). Outros dicionários etimológicos como os de Nascentes (1932, 1952) e Bueno (1963) não trazem informação precisa acerca do *terminus a quo* dos seus verbetes. Com vistas à retrodatação, o projeto DELPo (*Dicionário Etimológico da Língua Portuguesa*), vinculado ao NEHiLP (www.nehlp.org), desenvolveu um programa de computador chamado *Moedor*, por meio do qual simultaneamente (1) se verifica a inserção prévia da variante ortográfica do texto, apresentado sob a forma de edição diplomática; (2) se associa a um lema mediante um módulo de programação previamente alimentado, chamado *Desflexionador*; (3) atualiza a associação entre lemas e variantes no módulo supracitado mediante inserção feita por seus colaboradores; (3) sugere lemas a ser verificados pelos pesquisadores que colaboram para a confecção progressiva do dicionário. O resultado final dos muitos pesquisadores é disponibilizado *online* com devidos créditos aos seus colaboradores. Para tal objetivo, têm-se buscado as informações em obras ao mesmo tempo antigas e lexicalmente ricas. Dicionários preenchem esses dois requisitos. A obra *Hieronimi Cardosi Lamacensis dictionarium ex lusitanico in latinum sermonem* (1562-1563) é o mais antigo dicionário bilíngue da língua portuguesa

conservado até hoje, com 12069 entradas, nas quais se registra grande número de palavras existentes na segunda metade do século XVI, algumas das quais são comuns à língua portuguesa atual, enquanto outras são hoje inexistentes ou relegadas a variantes atuais menos conhecidas da língua portuguesa. Do mesmo autor é sua obra póstuma de 1599- 1570 *Dictionarium latinolusitanicum & vice versa*, preparado por Sebastião Stockammer. Da análise dessas obras, surgem as seguintes colaborações para a Linguística Românica e, mais especificamente, para a Etimologia de língua portuguesa: - a contribuição para o estudo do *terminus a quo*, ou primeira ocorrência do lema, em língua portuguesa, a qual pode ser retrodatada (é o caso de “cagalume”, “galacrista”, “malvasia”, “mocho”, “mofo”, “trevo”). Também se beneficia o estudo do *terminus ad quem* de alguns vocábulos e da reflexão sobre a última ocorrência da palavra afora de dicionários e poemas arcaizantes (por exemplo, a palavra “adem”). Muitas dessas palavras ainda hoje não se encontram dicionarizadas, a despeito de seu uso ainda ocorrer ou não (é o caso de “chisme”, “mugem”, “papalva”, “pelitre” e “ruiponto”). Tal questão se reflete posteriormente no abandono de uma das formas menos cultas (por exemplo, o de “alicornio” em prol de “unicórnio”, ambas atestadas no dicionário analisado, ou o de “alifante” que posteriormente será substituído por “elefante”) ou de especificações de uso (por exemplo, o verbete “alimária” será duplicado posteriormente por meio de uma forma mais culta “animal” simultaneamente ao fato de “alimária” se arcaizar ou se especializar em alguns tipos de discursos); - a datação de fenômenos fonéticos de adição e subtração fônica, sobretudo fatos considerados hoje como próteses, sínopes e epênteses (como em “abespa”, em “balea” e em “cangrejo”, atuais “vespa”, “baleia” e “caranguejo”). Também fenômenos de transposição nas chamadas metáteses (como em “briguigam”, atual “berbigão”); - a datação de variantes puramente gráficas, sem prejuízo da pronúncia sob a perspectiva da reconstrução fonética (por exemplo, “auestrux” em vez do atual “avestruz”, “codornis” e não “codorniz”). Os casos mais interessantes, para além da questão filológica da variação gráfica, são os que envolvem temas polêmicos como a suposta distinção fonológica apresentada pelos grafemas <ç> e <s>, que aparece questionada nas obras analisadas (por exemplo, ao lado de “passaro” há a grafia “paçaro”); - a datação de variantes não-gráficas do mesmo lema (por exemplo “ameigea” é a forma testemunhada, que difere da forma atual “amêijoa”, idem para “basalisco” e não, como hoje, “basilisco” e, por fim, “pintasirguo” para o atual “pintassilgo”). São de particular interesse casos de oscilação na mesma obra (como em “bibora” ou “vibora” ou então em “cirne” e “cisne”). Há muita confusão, tal como reflete a indicação lexicográfica abreviada como “fhist.” (isto é, “formas históricas”) no Grande Dicionário Houaiss, entre esse tipo de abonação (aparentada com a questão supracitada da datação dos fenômenos fonéticos por tratar-se de meras transformações fônicas a serem investigadas), do que com a penúltimo caso, em que apenas a grafia deveria ser levada em consideração. Nesses casos, algumas investigações são importantes, tais como a datação de um suposto alçamento de vogais pretônicas (como em “bisouro”, grafado hoje “besouro”, mas pronunciado em muitas variantes do português com uma vogal fechada e não com uma semifechada, idem para “curuja” e “gingibre” em vez das atuais “coruja” e “gingibre”); - a contribuição para a morfologia histórica, pois não só há formações hoje inexistentes, mas também possibilita o entendimento da mudança formal e semântica dos afixos (como em “rosairo”, atual “rosário”). A datação de variantes com afixos e afixoides variados que possam diferir das formas presentes, em lemas de obras posteriores (por exemplo, “cegarega”, alternado com “cigarra”) interessa a esta pesquisa, assim como os casos de *pluralia tantum* arcaicas que hoje se indicam no singular (como em “alforuas”, “armoles”, “mentrastos”, “bredos”, “cerralhas”). - a contribuição para questões de semântica histórica, seja na datação de diferentes acepções (por exemplo, a palavra “agulha” referente quer a um objeto, quer a uma espécie de peixe), seja na investigação de certos

hiperônimos usados lexicograficamente para esclarecer alguns verbetes por meio dos quais se transparecem algumas das classificações da época (como em “lula peixe” ou em “toninha peixe”, atuais seres classificados, respectivamente, como molusco e mamífero, ao mesmo tempo que “peixe” se usa também para outros zoônimos, a saber: “agulha”, “boga”, “carapao”, “congro”, “dourada”, “mugem”, “pargo”, “cabra”, “salmão” e “sarda”). O escopo de um dicionário etimológico alicerça-se em datas o mais precisas possível, que recuem os seus primeiros testemunhos, a fim de que se elaborem hipóteses linguísticas e etimológicas plausíveis. No Grande Dicionário Houaiss, palavras como “lêndea”, presentes em Cardoso, são vagamente apresentadas como sendo do século XVI. No verbete “azêmola”, o *terminus a quo* é associado ao século XIV embora os únicos testemunhos apresentados sejam *azemalla* e *azemella*, como vogal pós-tônica distinta. Assim sendo, a importância lexicográfica de Jerônimo Cardoso exerce-se tanto em questões metalinguísticas quanto na reconstrução de uma sincronia pretérita da língua portuguesa, a saber, a segunda metade do século XVI. Além de fornecer as primeiras abonações de muitos verbetes, retrodatando inclusive informações existentes nos dicionários etimológicos disponíveis, essa obra apresenta grande potencial para a comparação com outras línguas românicas coetâneas.