

## CERCETAREA LEXICOGRAFICĂ ACADEMICĂ ACTUALĂ: TRADIȚIE ȘI MODERNITATE

GABRIELA HAJA\*\*, ELENA ISABELLE TAMBA\*\*

1. Lucrarea de față propune un scurt istoric al activității Departamentului de lexicologie-lexicografie al Institutului de Filologie Română „A. Philippide”, Academia Română, Filiala din Iași, precum și descrierea strategiei proiectate și urmate de cercetătorii ieșeni în realizarea de programe, instrumente și resurse digitizate, adaptate normelor de elaborare a *Dicționarului limbii române* (DLR), din perspectiva principiilor coșeriene privitoare la epistemologia lingvisticii<sup>1</sup>. Considerăm necesară o atare abordare a subiectului în contextul actual, în care lexicografii din cele trei institute de profil ale Academiei Române implicate în redactarea *Dicționarului limbii române* (DLR) refac seria Pușcariu „după modelul DLR cu material nou, adus la zi, care să reflecte transformările care au avut loc în secolul scurs de la primele volume”<sup>2</sup>, cu scriere în format bază de date, conform standardelor internaționale de elaborare a dicționarelor.

2. Departamentul de lexicologie-lexicografie din Iași are o istorie de peste cincizeci de ani, perioadă în care mai multe generații de lexicografi s-au format și au contribuit la redactarea volumelor DLR asumate. Au fost realizate, astfel, opt volume, corespunzătoare literelor *E* (publicat în două părți, 2009, 2010), *L* (publicat în două

\* Abordări ale proiectelor de cercetare fundamentală.

\*\* Institutul de Filologie Română „Alexandru Philippide” al Academiei Române, Iași, str. Th. Codrescu, nr. 2, România.

<sup>1</sup> Este vorba despre principiile pe care Eugeniu Coșeriu le-a expus în mai multe rânduri, de la *Discursul de recepție* la Academia din Heidelberg și până la prezentarea făcută în cadrul sesiunii de prelegeri ținute în anul 1996 la Institutul de Fonetica și Dialectologie „Al. Rosetti”, București. Le vom regăsi definite și descrise în texte precum: Coșeriu 1978, Coșeriu 1992, Coșeriu 1993, Coșeriu 1994a, Coșeriu 1994b, Coșeriu 2004. Aceste principii sunt: 1. principiul obiectivității absolute; 2. principiul umanității sau al omului care, în lingvistică, este principiul vorbitorului; 3. principiul tradiției; 4. principiul antidogmatismului; 5. principiul responsabilității sociale.

<sup>2</sup> Acad. Marius Sala, prezentarea *Dicționarului-tezaur al limbii române*, reeditarea integrală, în 19 volume, a celor 37 de părți ale DLR, cu ocazia celei de a 145-a aniversări a Academiei Române, cf. „Academica”, nr. 4–5, aprilie–mai, anul XXI (246–247) ([http://acad.ro/academica2002/rev2011/pag\\_acad2011\\_nr246-247.htm](http://acad.ro/academica2002/rev2011/pag_acad2011_nr246-247.htm)).

tomuri, 2008), *Ș* (publicat într-o singură parte, 1978), *V, W, X, Y* (publicat în trei părți, 1997, 2002, 2005), care totalizează 3 950 de pagini ce cuprind 26 315 cuvinte și variante<sup>3</sup>. Dintre acestea, tomul VI, *Litera L (li-luzulă)* a fost finalizat în colaborare cu departamentul corespunzător de la Institutul de Lingvistică și Istorie Literară „Sextil Pușcariu” al Academiei Române, Filiala din Cluj.

Anul 2010 rămâne ca reper în istoria culturii române pentru încheierea primei ediții a DLR. În același an a fost concretizată și prima variantă a *Dicționarului limbii române* în format electronic, în cadrul proiectului de anvergură națională eDTLR<sup>4</sup>.

**2.1.** Istoria DLR a fost determinată de însăși istoria României moderne și contemporane, de istoria politică și de istoria culturală, pentru că proiectul *Dicționarului* Academiei s-a ivit în același timp cu formarea instituțiilor fundamentale ale statului modern, între care un loc aparte îl ocupă Academia Română. Redactarea primei ediții a fost grevată de două războaie mondiale și de o succesiune de regimuri politice: dacă politicile conservatoare de la începutul secolului al XX-lea puteau avea o influență discretă asupra selecției cuvintelor „cu drept de intrare” în *Dicționarul* Academiei, prin rezerva față de neologism, în aceeași perioadă, politicile culturale unioniste, de susținere, prin apelul la argumentul lingvistic, a unității naționale a românilor din Transilvania și a celor din Basarabia și Bucovina cu cei din Regatul României au făcut să fie tratat cu deosebită luare aminte lexicul popular și regional al dacoromânei, fără a fi neglijate referirile la dialectele române vorbite la sudul Dunării, relația dintre lexicografie și geografia lingvistică fiind considerată cu atenție de coordonatorii de până în prezent ai acestui proiect. Totodată, deschiderea europeană și atașamentul panromanice s-au manifestat, în volumele coordonate de Sextil Pușcariu, prin opțiunea de a oferi corespondentele franceze ale cuvintelor-titlu și ale principalelor sensuri în limba franceză, spre a facilita accesul la informația din dicționar pentru specialiștii străini. Schimbarea regimului politic după al Doilea Război Mondial a provocat între volumele publicate până în 1947 și cele ulterioare un hiatus, marcat de prezența în titlul volumelor tipărite după 1965 a sintagmei *s e r i e n o u ă*. Totuși, „cele două părți prezintă o anumită continuitate și formează un tot, asemănările putând fi considerate mai importante și mai numeroase decât deosebirile” (Sala *et al.* 2010: 7).

Din punctul de vedere al viziunii de alcătuire a listei de cuvinte, al tehnicii lexicografice, al consecvenței și omogenității redacționale, seria nouă prezintă progrese semnificative. Este avută în vedere, în continuare, reprezentarea obiectivă a identității și a specificului lingvistic din arealul dacoromânei, lexicul literar și cel popular ori regional bucurându-se de un tratament lexicografic egal, în virtutea faptului că DLR analizează deopotrivă varianta literară, cultă, și variantele

<sup>3</sup> *Dicționarul limbii române* cuprinde peste 175 000 de intrări și are 17 885 de pagini, publicate în 14 tomuri, apărute în 37 de volume. Seria nouă însumează 14 735 pagini.

<sup>4</sup> <https://metashare.info.uaic.ro/upload/eDTLR-UAIC.pdf>.

populară și regională de actualizare a limbii<sup>5</sup>. Faptul istoric al închiderii granițelor României și al escaladării politicii de izolare de după jumătatea secolului trecut are repercusiuni în decizia nefericită de a se renunța la echivalarea în franceză a lemelor și a sensurilor, precum și la citarea punctelor de anchetă dialectală din afara spațiului geopolitic al României socialiste, ceea ce a implicat, între altele, pierderea datelor privitoare la relația cu dialectele sud-dunărene.

**2.2.** Cu toate reproșurile ce i se pot aduce, DLR reprezintă o culme a creației lexicografice academice românești și este un reper fundamental pentru toți cercetătorii preocupați de studiul limbii și al culturii române<sup>6</sup>. Redactat după cele mai valoroase norme semasiologice și onomasiologice, DLR propune o lexematică implicită, căci, cu tot efortul pe care îl presupune analiza semnificantului și a semnificatului pe baza unui ansamblu de idiolecte<sup>7</sup> organizate cronologic, dicționarul înregistrează informații privind raporturile semantice complexe între cuvinte (cu precizarea datelor diatopice, diastratice și diafazice). De aceea, este posibilă identificarea de structuri lexematice paradigmatică și sintagmatică sau solidarități lexicale<sup>8</sup>, prin analiza fie a datelor prezente la nivelul definiției, fie a celor de la nivelul indicațiilor etimologice. Analiza este fiabilă fiindcă se pleacă de la limbajul realizat, adică de la vorbire, cu referire permanentă la normă și sistem, spre o

<sup>5</sup> De aceea ordinea citatelor ilustrative nu este întâmplătoare în succesiunea literar/popular, general/particular. Organizarea rigid cronologică, după criterii extralingvistice (anul scrierii sau al editării, anul revistei în care a apărut un text) nu oferă informații suficiente și absolute despre apariția ori dispariția unei forme ori a unui sens, deoarece indicarea anului este una relativă, dacă avem în vedere că, de exemplu, un autor ca Sadoveanu publică vreme de peste cinci decenii fără modificări esențiale de limbă și, mai ales, pentru că ansamblul izvoarelor textuale ale DLR nu constituie totalitatea textelor scrise în română. Cu toate acestea, există, în tradiția DLR, o normă potrivit căreia se precizează, în anumite cazuri, anul primei atestări înregistrate. Informația cronologică ar putea fi indicată, în volumele redactate în format electronic, la fiecare citat, așa cum propun colegii bucureșteni, dar nu în corpul articolului, ci ca metadată, alături de celelalte adnotări. Trebuie să facem observația că această propunere de adăugare a cronologiei exacte a atestărilor nu este un argument pentru schimbarea siglelor folosite în seria nouă a dicționarului și intrate în uzul general al specialiștilor. Modificarea generală a siglelor și a manierei de citare schimbă obiectivul inițial: acela al unificării și actualizării celor două serii ale lucrării.

<sup>6</sup> În prezent, DLR face parte, de exemplu, din bibliografia de referință a unor proiecte europene, precum *Dictionnaire Étymologique Roman* (DÉRom). Première phase: le noyau panroman, 2008–2014, proiect franco-german finanțat de Agence Nationale de la Recherche (ANR) și Deutsche Forschungsgemeinschaft (DFG); directori: Éva Buchi, Wolfgang Schweickard (<http://www.atilf.fr/DERom>) și *Lessico etimologico italiano* (LEI), proiect DFG, coordonat de Max Pfister și Wolfgang Schweickard, Saarbrücken (<http://www.accademiadellacrusca.it/en/accademia/connected-institutions-and-collaborations/lessico-etimologico-italiano>).

<sup>7</sup> Se poate discuta, de exemplu, faptul că *Bibliografia* DLR propune, ca izvoare directe, o serie de idiolecte sau limbaje individuale ori de limbaje speciale care actualizează sistemul limbii printr-o abatere mai mică ori mai mare de la norma observată în evoluție diacronică.

<sup>8</sup> Ne referim aici la teoria coșeriană a structurilor lexematice, așa cum a fost ea prezentată inițial în Coșeriu 1968: 3–16; textul a fost tradus din limba franceză de Silviu Berejan și publicat în Coșeriu 1992: 41–52; teoria a fost mult dezvoltată în prelegerile pe care savantul le-a susținut la diverse universități.

analiză lexematică, fie și parțială, făcută, implicit, prin intermediul prezenței arhilexemelor în definiții, al definirii prin sinonimie parțială și, explicit, prin corelațiile semantice rafinate, la nivel de sens și de subsens.

Ne apropiem astfel, pornind de la forma actuală a acestui dicționar, de o potențială extrapolare a unei lingvistici a vorbirii, fiindcă, deși cuvintele-titlu sunt tratate separat, în ordine alfabetică, ele sunt analizate semantic în funcție de context, așadar în funcție de raporturile semantice și formale dintre cuvinte. Această extrapolare este făcută empiric de orice lector specializat al dicționarului, însă ar putea fi concretizată, la nivelul întregului dicționar, și prin programe de analiză asistată de calculator a limbajului. Firește că nu se întrevede o analiză lexematică perfectă (aceasta trebuie, întotdeauna, verificată și validată de un lingvist) și completă (dicționarul nu cuprinde vocabularul limbii române în întregul său). Rezultatele ar putea fi însă cel puțin interesante, fiindcă s-ar putea reordona informațiile semantice din întregul dicționar, în structuri lexematice, pe baza raporturilor semantice primare (câmpuri și clase lexicale) și a celor secundare (modificare, dezvoltare, compoziție) existente în dicționar, cu implicații majore în utilizarea și valorificarea științifică a tezaurului limbii române. Pentru aceasta este necesară respectarea unor condiții *sine qua non* în procesul de unificare și actualizare a DLR, precum aceea a deplinei obiectivități științifice (aceasta înseamnă că datele lingvistice trebuie preluate corect și fidel cu originalul), aceea a respectării tradiției, a consecvenței și continuității, a păstrării echilibrului între analiza dimensiunilor limbajului (cronologie, spațiu, uz – cu apropiere ori abatere de la normă).

Aceste principii elementare în lingvistică se impun de la sine într-o lucrare lexicografică ce pornește de la limbaj spre teoria limbii, de la expresie spre normă și sistem, deoarece modificarea izvoarelor textuale ori schimbarea principiilor științifice de analiză a acestora poate genera concluzii eronate.

**3.** În continuare, vom prezenta etapele parcurse în cadrul Departamentului de lexicologie-lexicografie în direcția informatizării cercetării lexicografice fundamentale.

**3.1.** În ultimele decenii, cercetarea lingvistică (în general) și cea lexicografică (în special) din întreaga lume trece printr-un proces intens de schimbare, de adaptare la noile tehnologii, de modernizare a mijloacelor de redactare, de consultare etc., prin intermediul unor demersuri interdisciplinare. România s-a aliniat acestor tendințe, astfel încât în ultimii ani se poate vorbi despre un demers de informatizare a manierei de lucru în lexicografie. Acest fapt nu presupune o schimbare a principiilor lexicografice de realizare a dicționarelor, ci o actualizare a metodelor de documentare, a instrumentelor și a tehnicilor de lucru, a modalităților de difuzare a informației (Tamba *et al.* 2012: 259). În acest sens, în ultimii ani se poate observa conturarea unor strategii generale de informatizare a cercetării lingvistice, manifestate prin colaborări între institutele de filologie și cele de informatică ale Academiei Române, în parteneriat cu instituții de învățământ

superior, precum Facultatea de Informatică a Universității „Alexandru Ioan Cuza” din Iași.

O etapă specială și actuală în evoluția lexicografiei române o reprezintă informatizarea demersurilor lexicografice, ceea ce presupune, pe de o parte, digitizarea resurselor existente, iar pe de altă parte, crearea de dicționare direct în format electronic. Paralel, se continuă crearea de resurse lexicografice în format clasic<sup>9</sup>.

Premisa procesului de informatizare derulat în departamentul nostru, începând cu anul 2003, a fost cea care a stat și la baza strategiei de cercetare pe care am definit-o pe parcurs, anume că cercetările de lingvistică informatică și tehnologiile aferente se subsumează specificului lingvistic și lexicografic al DLR, într-un raport de determinare unidirecțională, ceea ce înseamnă că modificările formale și de conținut ale dicționarului nu trebuie să fie influențate de circumstanțe extralingvistice (incluzând aici și tehnologiile de prelucrare a limbajului ori de redactare adnotată<sup>10</sup> sau imitația unor modele, prestigioase, dar diferite ca structură și concepție lexicografică).

**3.2.** În scopul digitizării resurselor lexicografice existente, în ultimii ani s-a lucrat la informatizarea DLR, realizată prin proiectul complex *eDLR. Dicționarul-tezaur al limbii române în format electronic* (cu finanțare națională, pentru perioada 2007–2010), care a reunit un consorțiu format din informaticieni, specialiști în lingvistică computațională și prelucrarea limbajului natural, și lingviști lexicografi, ce aparțin mediilor academice și universitare din București, Iași și Cluj-Napoca<sup>11</sup>. Proiectul a urmărit mai ales transpunerea în format electronic a celor 17 885 de pagini ale *Dicționarului* (doar pentru ultimele volume tipărite existând o formă culeasă pe calculator), lucru care permite consultări complexe ale acestuia, precum și continuarea activităților de editare și de actualizare.

Informatizarea acestei importante lucrări are numeroase motive, între care posibilitatea de a o face cunoscută publicului larg (din cauza dimensiunilor mari, lucrarea se găsește în formă completă doar în institutele lingvistice ale Academiei

---

<sup>9</sup> Pentru o prezentare detaliată a istoriei lexicografiei românești, dar mai ales a procesului de informatizare a acesteia, vezi Tamba 2014 și Clim 2015.

<sup>10</sup> Prin *redactare adnotată* înțelegem redactarea cu mijloace informatice, în format bază de date, în care textul propriu-zis al dicționarului este completat, prin adnotare, cu metadate de diverse facturi și grade de finețe (de la indicarea cuvântului-titlu la analiza gramaticală a tuturor elementelor lexicale și până la posibile precizări de natură semantică, prin care se pot stabili relațiile de sub- și supraordonare, apartenența la un câmp sau la o clasă semantică, relațiile de solidaritate lingvistică etc.).

<sup>11</sup> Proiectul a fost condus de Facultatea de Informatică din cadrul Universității „Alexandru Ioan Cuza” din Iași, director: prof. dr. Dan Cristea. Din echipa de cercetare au făcut parte mai cu seamă specialiști lexicografi și informaticieni de la Academia Română: Institutul de Lingvistică „Iorgu Iordan – Al. Rosetti” din București, Institutul de Filologie Română „A. Philippide” din Iași și Institutul de Lingvistică și Istorie Literară „Sextil Pușcariu” din Cluj-Napoca, apoi Institutul de Informatică Teoretică din Iași și Institutul de Cercetări pentru Inteligență Artificială din București, precum și specialiști lingviști și informaticieni de la Universitatea „Alexandru Ioan Cuza” din Iași, Facultatea de Litere și Facultatea de Informatică (Cristea *et al.* 2007, Cristea, Haja 2011, Haja 2011).

și în câteva biblioteci), necesitatea actualizării și unificării (ca urmare a diferențelor mari de timp, apar discrepanțe, la nivelul perspectivei de stabilire a listei de cuvinte, al redactării lexicografice, al bibliografiei DLR între literele lucrute în prima serie și cele din ultimele volume), ca și relevanța utilizării sale în cercetări de analiză asistată de calculator a limbajului (de exemplu, numărul semnificativ de citate care însoțește fiecare sens al unui cuvânt constituie un valoros corpus adnotat semantic pe care pot fi antrenate programe de dezambiguizare).

Pornită ca ipoteză, posibilitatea creării programelor și a instrumentelor informatice necesare transpunerii și, ulterior, redactării DLR în format digitizat a fost verificată în cadrul cercetărilor întreprinse de noi în perioada 2003–2005 (Haja *et al.* 2005). Rezultatele obținute atunci au fost confirmate în proiectele specifice, derulate ulterior (Haja *et al.* 2008). Proiectul eDLR a fost anunțat și pregătit de o serie de alte granturi<sup>12</sup> care au demonstrat fezabilitatea ideii și au creat unele instrumente necesare pentru realizarea sa.

**3.2.1. Dicționarul limbii române (DLR) în format electronic. Studii privind achiziționarea,** grant cu finanțare națională, desfășurat în perioada 2003–2005 la Institutul de Filologie Română „A. Philippide”, Iași (director de proiect: dr. Cristina Florescu). Prin acest proiect s-a verificat și demonstrat posibilitatea transformării *Dicționarului limbii române* din text tipărit în text electronic adnotat, prelucrat cu ajutorul unui program specific, DLReX – un instrument de achiziționare, prelucrare și consultare a DLR, bazat pe o euristică prin care sunt recunoscute diferitele câmpuri formale ale textului unui articol, putându-se identifica automat textul definițiilor, al citatelor sau al siglelor<sup>13</sup>.

**3.2.2. Resurse lingvistice în format electronic: Monumenta linguae Dacoromanorum. Biblia 1688. Regum I, Regum II – ediție critică și corpus adnotat. (MLD. Biblia 1688),** grant cu finanțare națională, desfășurat în perioada 2006–2007 la Institutul de Filologie Română „A. Philippide”, Iași (director de proiect: dr. Gabriela Haja) – proiect prin care a fost găsită o posibilă metodă de achiziționare în format electronic a unor izvoare de limbă română veche din *Bibliografia DLR*, cu aplicație asupra a două cărți din *Biblia* tipărită la București în anul 1688 (*A împărățiilor cea dentâiu*, *A împărățiilor a doua*), și au fost create instrumente de indexare și adnotare automată (analiză morfo-sintactică), la nivel de cuvânt, a textelor în limba română veche.

**3.2.3. DLRI. Bază lexicală informatizată. Derivate,** grant cu finanțare națională, desfășurat în perioada 2007–2008 la Institutul de Filologie Română „A. Philippide”, Iași (director de proiect: dr. Cristina Florescu). În cadrul acestui

<sup>12</sup> Un alt proiect care a avut ca scop digitizarea unei resurse lexicografice importante, reper al lexicografiei române moderne, este *Lexiconul de la Buda (1825). Ediție emendată și prelucrată electronic pentru consultare on-line* (desfășurat la Universitatea „Babeș-Bolyai”, Cluj-Napoca, pentru perioada 2011–2013; director de proiect: dr. Maria Aldea). Rezultatele proiectului și ediția electronică pot fi consultate la adresa <http://www.bccluj.ro/lexiconuldelabuda>

<sup>13</sup> Pentru detalii legate de acest proiect, vezi și Haja *et al.* 2005.

proiect s-a realizat un eșantion lexicografic format din derivatele pe terenul limbii române cu sufixul *-ime* de origine latină, și cele cu *-iște* – de origine veche slavă, din seria veche DA și din seria nouă a DLR, precum și unificarea tehnico-lexicografică a articolelor DA – DLR, prin mijloace informatice.

**3.2.4.** Cercetările din cadrul eDTLR au evidențiat ideea necesității unui corpus lexicografic românesc<sup>14</sup>, realizată prin proiectul CLRE. *Corpus lexicografic românesc esențial. 100 de dicționare din bibliografia DLR aliniate la nivel de intrare și la nivel de sens* (finanțat de CNCS-UEFISCDI pentru perioada 2010–2013), care s-a desfășurat în cadrul departamentului nostru (director de proiect: dr. Elena Tamba). Acest proiect a avut ca scop crearea unui alt instrument de lucru foarte necesar lexicografilor și util tuturor celor interesați de studiul limbii române. Astfel, cercetătorii implicați în proiect au realizat o bază de date care cuprinde dicționare esențiale din *Bibliografia DLR*, aliniate la nivel de intrare, au construit un mediu de programe care permite consultarea interactivă a acestui corpus, care este pe cale să devină un cadru modern de cercetare lexicografică, ușor adaptabil la o diversitate de obiective științifice; de asemenea, a fost generată o listă cvasiexhaustivă de cuvinte ale limbii române, pornind de la corpusul aliniat<sup>15</sup>. Proiectul a presupus, pe de o parte, alinierea tuturor cuvintelor din dicționare și, pe de altă parte, alinierea acestora cu intrările din eDTLR. Consultarea informațiilor din corpus se poate face fie la nivelul unui singur dicționar, fie la nivelul întregii colecții, permițându-se căutarea după litere sau după intervale (expresii regulate). Deocamdată, corpusul poate fi consultat în scop de cercetare, pe bază de acces securizat (cu ID personalizat și parolă), la adresa: <http://www.lexi.philippide.ro/clre>.

**3.3.** Din anul 2014, proiectul CLRE a fost inclus ca o nouă direcție de cercetare în planul Academiei Române, ca activitate de bază a lexicografilor din departamentul ieșean de profil, alături de redactarea DLR, *Litera C*.

Proiectul CLRE permite stabilirea de legături cu alte proiecte naționale (eDTLR și TDRG<sup>316</sup>) și internaționale (fie prin participarea unor lexicografi ieșeni în proiecte lexicografice internaționale<sup>17</sup>, fie prin includerea unor proiecte de lexicografie a limbii române în rețele lexicografice majore internaționale<sup>18</sup>).

<sup>14</sup> Pentru o prezentare detaliată a acestor aspecte, vezi Dănilă 2010.

<sup>15</sup> Pentru detalii legate de proiectul CLRE, vezi și Dănilă 2013; Tamba Dănilă *et al.* 2012.

<sup>16</sup> H. Tiktin, *Rumänisch-deutsches Wörterbuch*, 3., überarbeitete und ergänzte Auflage von Paul Miron und Elsa Lüder. Band I–III. Cluj-Napoca, Editura Clusium. I: 2000; II: 2003; III: 2005 în format electronic ([www.lexi.philippide.ro/tiktin](http://www.lexi.philippide.ro/tiktin)).

<sup>17</sup> Vezi, de exemplu, participarea câtorva lexicografi români la realizarea unor lucrări precum *Rumänisch-deutsches Wörterbuch* de Heimann Tiktin, ediția a doua – proiect internațional interuniversitar, Freiburg – Iași, 1979–1987; coordonatori: Paul Miron, V. Arvinte sau *Dictionnaire Étymologique Roman* (DÉRom) ș.a.

<sup>18</sup> Menționăm aici includerea corpusului lexicografic CLRE în programul *ENeL: European Network of e-Lexicography* (proiect internațional COST Action IS 1305, <http://www.elexicography.eu/>, acțiune în cadrul liniei de finanțare COST, prin care se urmărește crearea unei rețele pan-europene de resurse lexicografice atât pentru uzul academic, cât și pentru publicul larg). Proiectul se desfășoară pe durata a patru ani (2013–2017) și, până în prezent, sunt implicate 30 țări din Europa, precum și statul Israel.

După încheierea primei ediții a DLR, s-a trecut la reluarea și actualizarea, cu instrumente și tehnici de lucru moderne, a primei părți a *Dicționarului limbii române* (DA), publicat în intervalul 1913–1948, în prezent derulându-se activitățile de documentare și de redactare. Deocamdată, pentru redactarea articolelor DLR avem la dispoziție o colecție de izvoare textuale pentru limba literară și pentru limba populară ori regională (lucrările din *Bibliografia* DLR, pe suport de hârtie și în format .doc ori .txt obținut prin scanare și recunoaștere optică a caracterelor), dicționare, fișe (scrise de mână, dactilografiate sau în format electronic), un program de indexare și de căutare automată accesibil on-line. Pentru optimizarea procesului de documentare și de redactare a dicționarului-tezaur ar fi necesar un corpus diacronic al limbii române, adnotat la nivel de cuvânt, cu posibilități de consultare complexă (așa cum permite, de exemplu, *Sketch Engine*<sup>19</sup>), de programe de scriere de dicționar adaptate la specificul și complexitatea DLR<sup>20</sup> și, mai ales, de lexicografi specializați și în domeniul lexicografiei informatizate. Cercetătorii ieșeni s-au format și documentat în ultimii ani în această direcție, prin participarea la proiectele prezentate (prin lucrul direct, prin participări la școli de vară<sup>21</sup>, la *workshop*-uri<sup>22</sup> sau la manifestări științifice internaționale dedicate cercetărilor actuale din domeniu). O altă modalitate de informare a fost colaborarea cu colegii informaticieni în cadrul unor proiecte sau al întâlnirilor de lucru ale grupului „Natural Language Processing” de la Facultatea de Informatică a Universității „Alexandru Ioan Cuza” din Iași.

Unii dintre lexicografii ieșeni s-au specializat și prin studii doctorale și postdoctorale cu teme legate de lexicologie-lexicografie sau de informatizare a dicționarelor.

**3.4. Direcțiile de cercetare în lexicografia europeană actuală pot fi grupate astfel:**

- continuarea editării (cu termen de finalizare de 10–15 ani) a unor ediții clasice (vezi DWB – *Deutsches Wörterbuch von Jacob Grimm und Wilhelm Grimm*);
- realizarea variantelor electronice, accesibile on-line, ale unor dicționare, plecând de la corpusuri de texte din ce în ce mai mari și folosind instrumente variate (eventual, programe de extragere de exemple, după criterii specifice);

<sup>19</sup> Pentru informații privitoare la *Sketch Engine* (sistem de căutare în corpus, utilizabil ca instrument de lucru pentru lexicografi), vezi Dănilă *et al.* 2011: 194–195.

<sup>20</sup> Institutul de Lingvistică „Iorgu Iordan – Al. Rosetti” din București a inițiat o colaborare cu firma Oxygen, specializată și în realizarea unor interfețe dedicate scrierii de dicționare și care pregătește o variantă adaptată la normele de redactare DLR. S-au făcut și la Iași cercetări de acest tip în perioada 2010–2012, ajungându-se la o primă variantă de interfață dedicată redactării DLR, prezentată în cadrul *workshop*-ului „Lexicografia academică românească. Provocările informatizării”, organizat anual de departamentul nostru, și la alte manifestări științifice organizate în țară.

<sup>21</sup> De exemplu, „A doua școală de vară franco-germană în etimologie romanică”, ATILF, Nancy, organizată în cadrul priectului internațional DÉRom.

<sup>22</sup> A XI-a ediție a *workshop*-ului în lexicografie și lexicologie computațională „LEXICOM 2011”, Sankt-Petersburg, Rusia, 14–18 iunie 2011, organizat de ABBYY InfoPoisk LLC, Lexical Computing Ltd. și de Universitatea de Stat din Sankt-Petersburg.

- realizarea de corpusuri lexicografice;
- utilizarea sistemelor de scriere de dicționar (DWS)<sup>23</sup>, care permit alinierea ulterioară mai facilă a diverselor lucrări lexicografice, operație realizabilă pe baza unor protocoale de colaborare încheiate între diverse instituții cu profil lexicografic, editori etc.;

- crearea de rețele lexicografice internaționale.

Prin raportare la aceste tendințe din lexicografia europeană, în lexicografia română se are în vedere:

- continuarea (re)editării unor ediții clasice (DLR, DEX etc.) în formă tipărită și în variantă electronică;

- realizarea unor corpusuri de texte pentru limba română (în lucru);

- dezvoltarea de corpusuri lexicografice (în lucru – de exemplu, CLRE, prezentat *supra*);

- utilizarea programelor informatice de scriere de dicționar (în lucru) (Tamba *et al.* 2012: 266).

4. Așadar, după o studiere a posibilităților de transpunere în format electronic a DLR, s-a trecut la următoarea etapă de lucru, constând în crearea unor instrumente de analiză și adnotare a limbii române vechi, definindu-se și un model de format electronic pentru editarea electronică a volumelor din *Bibliografia DLR*, realizarea unui eșantion de articole unificate și actualizate din seria veche potrivit normelor DLR cu ajutorul instrumentelor informatice în curs de elaborare în departament, realizarea primei variante integrale a DA și DLR în format electronic, constituirea unui corpus diacronic de dicționare ale limbii române, aliniate la nivel de intrare și, în fine, integrarea cercetării lexicografice de nivel academic în circuitul internațional.

Putem observa că, urmând această strategie, cercetătorii ieșeni formați ca redactori ai DLR au avut, în ultimul deceniu, contribuții însemnate la procesul de adaptare a lexicografiei academice românești potrivit schimbărilor pe care le implică modernizarea instrumentelor de lucru în acest domeniu, la nivel internațional, făcând, în același timp, efortul de a adecva aceste instrumente și resurse, create sau adaptate pentru limba română, la rigorile și normele DLR.

Activitățile de cercetare din proiectele de informatizare derulate în departamentul nostru au conturat premisele punerii la dispoziția cercetătorilor români și străini a unor resurse lexicografice accesibile on-line, mult timp așteptate de specialiști, necesare studierii și promovării limbii și culturii române.

#### BIBLIOGRAFIE

Clim 2015 = Marius-Radu Clim, *La lexicografía rumana informatizada: tendencias, obstáculos y logros*, în vol. *Lexicografía de las lenguas románicas. Aproximaciones a la lexicografía moderna y contrastiva*, vol. II, coords.: María Dolores Sánchez Palomino y María José

<sup>23</sup> Pentru o analiză a sistemelor de scriere de dicționar, vezi Dănilă *et al.* 2011: 195–196.

- Domínguez Vázquez, eds.: María José Domínguez Vázquez, Xavier Gómez Guinovart y Carlos Valcárcel Riveiro, de Gruyter Verlag, p. 95–110.
- Coseriu 1968 = Eugenio Coseriu, *Les structures lexématiques*, în „Zeitschrift für französische Sprache und Literatur” (Wiesbaden), Beiheft, Neue Folge, Heft I, p. 3–16.
- Coseriu 1978 = Eugenio Coseriu, *Antrittsrede an der Heidelberger Akademie der Wissenschaften*, în „Jahrbuch der Heidelberger Akademie der Wissenschaften für 1977”, Heidelberg, p. 107–110.
- Coseriu 1992 = Eugenio Coseriu, *Principiile lingvisticii ca știință a culturii*, în „Apostrof” (Cluj-Napoca), III, nr. 11, p. 11, 14.
- Coseriu 1993 = Eugenio Coseriu, *Discurso pronunciado con motivo de su investidura como doctor honoris causa*, în „Discursos pronunciado en el acto de investidura de Doctor Honoris Causa del Excelentísimo Señor Eugenio Coseriu”, Granada, p. 21–35.
- Coșeriu 1992 = Eugeniu Coșeriu, *Structurile lexematice*, în „Revistă de lingvistică și știință literară”, (Academia de Științe a Republicii Moldova, Institutul de Filologie), nr. 6, p. 41–52.
- Coșeriu 1994a = Eugeniu Coșeriu, *Deontologia și etica limbajului*, în idem, *Prelegeri și conferințe (1992–1993)*, Iași, supliment al publicației „Anuar de lingvistică și istorie literară”, XXXIII, 1992–1993. Seria A. Lingvistică, p. 163–172.
- Coșeriu 1994b = Eugeniu Coșeriu, *Deontologia culturii*, în *ibidem*, p. 173–180.
- Coșeriu 2004 = Eugeniu Coșeriu, *Epistemologia lingvisticii*, în *In memoriam Eugeniu Coșeriu*, extras din publicația „Fonetica și dialectologie”, XX–XXI, 2001–2002, București, Editura Academiei Române, p. 87–94.
- Cristea *et al.* 2007 = Dan Cristea, Marius Răschip, Corina Forăscu, Gabriela Haja, Cristina Florescu, Elena Dănilă, Bogdan-Mihai Aldea, *The Digital Form of Thesaurus Dictionary of the Romanian Language*, în *Advanced in Spoken Language Technology*, Corneliu Burileanu, Horia-Nicolai Teodorescu (eds.), București, Editura Academiei Române, p. 195–206.
- Cristea, Haja 2011 = Dan Cristea, Gabriela Haja, *The Thesaurus Dictionary of Romanian Language in Electronic Form*, în „Clarin”, Newsletter of Clarin Project, nr. 13, January–June, p. 10–11, <http://www.clarin.eu/newsletter>.
- Dănilă 2010 = Elena Dănilă, *Despre necesitatea realizării unui corpus lexicografic românesc esențial*, în „Philologica Jassyensia”, anul VI, nr. 2 (12), p. 41–50 (și on-line, la adresa [http://www.philologica-jassyensia.ro/upload/VI\\_2\\_Danila.pdf](http://www.philologica-jassyensia.ro/upload/VI_2_Danila.pdf)).
- Dănilă 2013 = Elena Dănilă, *Corpus lexicographique roumain essentiel. Les dictionnaires de la langue roumaine alignés au niveau de l'entrée*, în *Actas del XXVI Congreso Internacional de Lingüística y de Filología Románicas 6–11 septiembre 2010, Valencia*, Casanova Herrero, Cesareo Calvo Rigual (eds.), de Gruyter Verlag, vol. VIII, p. 125–134.
- Dănilă *et al.* 2011 = Elena Dănilă, Marius-Radu Clim, Ana Catană-Spenchiu, *Towards a Romanian Lexicographic Corpus*, în „Philologica Jassyensia”, anul VII, nr. 2 (14), p. 191–198.
- Haja 2011 = Gabriela Haja, *Dicționarul limbii române în format electronic. Instrument pentru cercetarea și învățarea limbii române*, în *Simpozionul internațional Grigore Bostan – 70. Probleme actuale de filologie română*, Cernăuți, Zelena Bucovina, Herța, Grupul editorial „Cuvântul”, p. 221–230.
- Haja *et al.* 2005 = Gabriela Haja, Elena Dănilă, Corina Forăscu, Bogdan-Mihai Aldea, *Dicționarul limbii române (DLR) în format electronic. Studii privind achiziționarea*, Iași, Editura Alfa.
- Haja *et al.* 2008 = Gabriela Haja, Elena Dănilă, Marius-Radu Clim, *Premise ale informatizării cercetării lexicografice academice românești*, în *Limba română. Dinamica limbii, dinamica interpretării*, Editura Universității din București, p. 585–591.
- Sala *et al.* 2010 = Acad. Marius Sala, acad. Gheorghe Mihăilă, dr. Monica Busuioc, *Cuvânt-înainte la Dicționarul limbii române*, tomul I, A–B, București, Editura Academiei Române, p. 5–11.
- Tamba 2014 = Elena Tamba, *La lexicografía Rumana. Historia y Actualidad*, în *Lexicografía de las lenguas románicas. Perspectiva histórica*, eds. Félix Córdoba Rodríguez, Ernesto González Seoane, María Dolores Sánchez Palomino, vol. I, de Gruyter Verlag, p. 265–282.
- Tamba Dănilă *et al.* 2012 = Elena Tamba Dănilă, Marius-Radu Clim, Ana Catană-Spenchiu, Mădălin Patrașcu, *The Evolution of the Romanian Digitalized Lexicography. The Essential Romanian Lexicographic Corpus*, în *Proceedings of the 15 th EURALEX International Congress*, 7–11

august 2012, eds. Ruth Vatvedt Fjeld, Julie Matilde Torjusen, Oslo, Press Representrales, UiO, p. 1014–1017 (și on-line, la adresa: [http://www.euralex.org/proceedings-toc/euralex\\_2012/](http://www.euralex.org/proceedings-toc/euralex_2012/)).

Tamba *et al.* 2012 = Elena Tamba, Marius-Radu Clim, Ana Catană-Spenchiu, Mădălin Patrașcu, *Situația lexicografiei românești în context european*, în „Philologica Jassyensia”, anul VIII, nr. 2 (16), p. 259–268 (și on-line, la adresa: [http://www.philologica-jassyensia.ro/upload/VIII\\_2\\_Tamba\\_Clim.pdf](http://www.philologica-jassyensia.ro/upload/VIII_2_Tamba_Clim.pdf)).

### CONTEMPORARY ACADEMIC LEXICOGRAPHICAL RESEARCH: TRADITION AND MODERNITY

#### ABSTRACT

The paper presents a short history of academic researches within The Department of Lexicology and Lexicography from the “A. Philippide” Institute of Romanian Philology, Romanian Academy, Iași Branch, the linguistic and lexicographical principles concerning the bringing up to date of the first volumes of DLR. *Romanian Language Dictionary* (1<sup>st</sup> ed. 1913–2010), work in progress in Romanian institutes involved in the accomplishment of DLR, arguing the need to comply with the current editing rules of these outstanding lexicographical works. The argumentation is founded on the principles of the linguistic epistemology, in terms of the integral linguistics defined by Eugeniu Coșeriu. These principles can be identified in our strategy for the digitization process of the Romanian lexicographical research. The achievements in specific projects developed in our department are also described in detail in the second part of the paper.

**Keywords:** *linguistic & lexicographical principles, the history of the Romanian Language Dictionary, digitization process phases, tradition & innovation, linked data.*

