

Derivate în DTLR. Bază lexicală informatizată

Cristina FLORESCU, Elena DĂNILĂ, Laura MANEA,
Marius-Radu CLIM, Marius RĂSCHIP

1. Introducere

Comunicarea de față se referă la o serie din rezultatele cercetării cuprinse în proiectul *DLRI. Bază lexicală informatizată. Derivate* (cod CNC SIS 1609), desfășurat în perioada 2007-2008.

Analiza lingvistică și informatică a fenomenului derivării cu sufixele *-ime* și *-iște* pe teren românesc pleacă de la achiziționarea electronică a acestor derivate, realizându-se, astfel, în lexicografia românească, primul eșantion semnificativ care include un corpus de articole DA+DLR informatizat, unificat și actualizat. Proiectul, condus de CS I Cristina Florescu, se derulează în Institutul de Filologie Română „A. Philippide” al Academiei Române – Filiala Iași, iar echipa de cercetare este formată din lingviștii Laura Manea, Elena Dănilă, Marius-Radu Clim și informaticienii Bogdan Aldea, Corina Forăscu și Marius Răschip¹.

În ultima perioadă, în filologia românească s-au inițiat proiecte menite să ducă la unificarea, informatizarea, actualizarea lexicologică și bibliografică a lucrării fundamentale a lexicografiei românești, *Dicționarul Tezaur al Limbii Române* (DA+DLR).

Proiectul prezent își propune trei obiective.

1. Analizarea lingvistico-lexicografică și prelucrarea informatică a unui grup lexical format din toate derivatele pe terenul limbii române cu sufixul *-ime* (de origine latină) și cele cu *-iște* (de origine veche slavă) existente în DA, DLR, MDA și în bibliografia de strictă specialitate. Grupul lexical este reprezentat de 720 de articole lexicografice prelucrate în format XML – inițial cu ajutorul DLReX², corpus concretizat într-o bază de date lexicale informatizată cuprinzând aceste derivate. În funcție de materialul achiziționat în format electronic s-a realizat rafinarea instrumentului de lucru nou creat.

2. Unificarea lingvistică, tehnico-lexicografică a articolelor DA+DLR conform normelor DLR.

3. Redactarea unui volum care să cuprindă:

a) un dicționar tipărit însumând derivatele în discuție, înregistrat pe suport electronic și introdus on-line;

¹ Mulțumim și pe această cale d-lui prof. univ. Dan Cristea pentru sprijinul constant acordat.

² DLReX este un instrument de achiziționare, prelucrare și consultare a DLR, creat în cadrul grantului *Dicționarul limbii române în format electronic. Studii privind achiziționarea* (cod CNC SIS 1815), proiect desfășurat în perioada 2003 – 2005.

b) un studiu care să însumeze cercetările întreprinse de grupul de cercetători pe parcursul celor doi ani de proiect, în domeniile: lexicologie, istorie a limbii române și istorie a limbii literare, gramatică, lexicografie informatizată.

În cadrul cercetării de față se întreprinde pentru prima dată achiziționarea electronică a unor texte lexicografice grupate categorial atât din seria veche DA a *Dicționarului limbii române*, cât și din seria nouă DLR. În egală măsură, analiza vizează actualizarea și unificarea, cu ajutorul instrumentelor și tehnologiei create, a unui grup lexical semnificativ pentru limba română. Subliniem că, înaintea cercetării de față, respectivul grup lexical cuprindea articole lexicografice inegale în ceea ce privește informația lingvistică, redactate conform principiilor lexicografice vechi din DA și respectiv tehnicii noi din DLR.

Au fost realizate fazele prevăzute în proiect.

a) S-a stabilit lista de intrări. Precizăm faptul că cercetarea a plecat de la un număr de cca 300 lexeme, stabilite în funcție de lucrările de specialitate anterioare, ajungându-se la o cifră de 720 de lexeme înregistrate în bibliografia de bază.

b) S-a făcut scanarea articolelor care cuprind derivatele respective din DA și din DLR (o parte din materialul scanat a fost selectat din corpul de date lexicografice DA+DLR din proiectul complex eDTR³).

c) S-a realizat OCR-izarea⁴ materialului rezultat.

d) S-a finalizat corectarea materialului OCR-izat. Precizăm că, dacă pentru materialul DLR – în cadrul grantului *Dicționarul limbii române în format electronic. Studii privind achiziționarea* – operația de OCR-izare mai fusese întreprinsă, prin cercetarea de față, pentru prima oară este făcută OCR-izare unor articole, grupate lexicologic, din DA.

e) Operația de maximă dificultate a fost reprezentată de redactarea, în funcție de normele DLR, a unui număr de cca 220 articole cu derivate în *-ime* și *-iște* din porțiunea cuprinsă în DA; respectiva operație a presupus documentare lingvistică, fișare, stabilire a definițiilor, structurare a arborelui lexical, propunere a unor noi etimologii.

f) Este în curs de desfășurare rafinarea, în funcție de materialul lexicografic, a instrumentului informatic de lucru.

g) Este pe cale de finalizare prelucrarea lingvistico-semantică a corpului de date astfel obținut.

2. Analiza computațională

Baza lexicală informatizată presupune realizarea unei interfețe în care să se poată face consultarea întregului corpus de derivate obținut în cadrul proiectului.

În urma operațiilor menționate anterior a rezultat o listă de cuvinte care a fost prelucrată informatic după cum urmează:

a) lista editată în Microsoft Word a fost salvată în formatul Web Page Filtered „.html”;

³ Proiect complex eDTR – *Dicționarul Tezaur al Limbii Române în format electronic* (2007–2010).

⁴ Transpunerea din format imagine (.tif) în format text (.rtf).

- b) formatul rezultat „.html” a fost transformat în așa fel încât să respecte standardul XHTML, folosind parserul Mozilla;
- c) în formatul transformat, de tip XML, intrările de dicționar au fost identificate și salvate în fișiere (pagini XHTML) separate;
- d) referințele („v.” intrare) au fost transformate în legături dintre pagini;
- e) în final s-a creat un index alfabetic.

Transformarea în format XML și parsarea sumară a intrărilor de dicționar au fost implementate în limbajul de programare Java, folosindu-se diverse biblioteci cum ar fi: dom4j, htmlparser, mozillahtmlparser.

Păstrarea fiecărei intrări de dicționar într-un fișier separat aduce mai multă flexibilitate în exploatarea colecției. Într-o abordare simplă, aceste fișiere pot fi indexate cu un motor de căutare on-line. Cuvântul titlu al intrării este și titlu de pagină, iar acest lucru crește importanța paginii la o eventuală căutare a lui, făcând ca un motor de căutare să schimbe ordinea rezultatelor în funcție și de preferințele exprimate de utilizatori. O pagină aleasă mai des va primi o poziție mai importantă.

3. Analiză lingvistică

Din punct de vedere lingvistic, în proiectul de față se urmăresc o serie de relevanțe specifice.

În primul rând, este urmărită evoluția fenomenului lingvistic din momentul ultimelor studii de specialitate din 1962 (anul celor din urmă lucrări focalizate asupra elementelor în discuție) și până astăzi, marcându-se importanța unei asemenea analize centrate asupra materialului exhaustiv din DA și DLR. Se studiază relevanța privind acest tip de analize valorificatoare a bogatului material lexicologic cuprins în DA+DLR.

Atragem atenția asupra faptului că ne confruntăm adesea cu păreri conform cărora se consideră că materialul lexical cuprins în DA+DLR este în totalitate cunoscut și analizat în literatura de specialitate, discuțiile analizelor lingvistice axându-se, mai totdeauna, asupra limitelor inerente acestui lexicon tezaur. Fără discuție, în mod normal și obligatoriu, DA+DLR cuprinde un material mereu depășit, pe măsura evoluției limbii și avansării cercetării lingvistice. Lexiconul în discuție se structurează ca un punct de plecare (preluat, completat și modificat) pentru cele mai multe analize ale cercetărilor lexicale contemporane. Finalizarea DLR creează posibilitatea de a avea o evidență, extrem de minuțioasă lexicografic, asupra unei întregi suite de fapte lingvistice importante care pot fi vizualizate pentru prima oară la nivel macro-lexicografic.

Discutând, punctual, cazul derivatelor din limba română, modificarea cantitativă a numărului derivatelor analizate până acum impune o reajustare a gândirii fenomenului. Altfel spus, într-un fel se prezintă generalizarea analitică (etimologică, semantică, gramaticală etc.) în cazul unui material lexical care cuprindea un corpus de cca 300 de lexeme, și cu totul altfel analiza unui corpus care conține în prezent 720 de lexeme. Parametrii analizei se amplifică, modificându-se semnificativ.

Asemenea tip de analize generalizatoare, care să însumeze *toate* datele lexicale cuprinse în corpusul DA+DLR = DTLR pot produce modificări asupra obiectului de studiu mult mai mult decât se consideră îndeobște.

În cazul de față, analiza lingvistică a lexemelor avute în vedere este întreprinsă din punctul de vedere al etimologiei, al structurii semantice și noționale, al categoriei gramaticale, al repartiției dialectale, stilistice etc.

Se înțelege că cercetarea valorifică toate analizele lingvistice anterioare care studiază aspecte semnificative privind formarea și structura derivatelor în limba română (Pascu 1916, Pașca 1948, Carabulea 1959, Sădeanu 1962).

Nu ne putem permite aici o descriere minuțioasă a rezultatelor lingvistice ale cercetării întreprinse. Ne vom opri la câteva aspecte semnificative în sensul celor spuse până acum.

Sub aspect derivativ, sufixele lexicale în discuție au un potențial productiv mare.

Într-o primă fază, proiectul nostru a pornit – pe baza studiilor de specialitate care identificaseră derivatele formate cu aceste sufixe – de la un număr de cca 180 derivate cu sufixul *-ime* și un număr de 140 derivate cu sufixul *-iște*, deci un total de cca 320 derivate cu aceste sufixe.

În urma stabilirii listei de cuvinte pe baza fișării DTLR și a lucrărilor de specialitate s-a ajuns la rezultate mult îmbunătățite: 720 derivate (din care cca 160 de intrări sunt variante). Dintre acestea cca 490 sunt derivate cu sufixul *-ime* și cca 230 cu sufixul *-iște*.

3.1. Sufixul *-ime*

Avem în vedere situația specifică pentru limba română a derivatelor cu sufixul *-ime* (< lat. *-imen*).

Din punct de vedere statistic, în română, sufixul *-ime* se atașează unor cuvinte de origine diversă (latină, slavă, turcă, maghiară etc.)⁵, remarcându-se o situație specială pentru limba română veche (unde sufixul *-ime* formează derivate substantive abstracte – de tipul *cruzime* < *crud*, *înălțime* < *înalt* etc.), față de limba română contemporană (unde sufixul *-ime* a primit mai ales o valoare colectivă – *românime* < *român*) (Iordan 1956: 311; Ivănescu 2000: 701).

Din punct de vedere semantic, sufixul *-ime* formează:

a) substantive feminine care exprimă:

– o dimensiune: *adâncime*, *afunzime*, *grosime*, *înălțime*, *îngustime*, *întinzime*, *lărgime*, *lătățime*, *lățime*, *lungime*, *mărime*, *micime*, *micșorime*, *mititecime*, *profunzime*, *scurtime*, *strâmtime*, *subțirime* etc.;

– o calitate: *acrimă*, *amărime*, *ascuțime*, *cruzime*, *frăgezime*, *grăsime*, *iscusime*, *iuțime*, *limpezime*, *netezime*, *prospețime*, *surzime*, *usturime*, *ușurime* etc.;

– o cantitate: *desime*, *mulțime*, *puținime*, *puținime*, *rărime* etc.;

– valoare temporală: *învechime*, *vechime* etc.

b) numerale fracționare: *cincime*, *cincizecime*, *doime*, *miime*, *milionime*, *treime*, *sutime*, *zecime* etc.;

⁵ Vezi și Carabulea 1959: 67.

c) substantive cu valoare colectivă, derivatele nou formate exprimând:

- o mulțime de oameni: *arăbime*, *arăpime*, *argățime*, *aristocrațime*, *armenime*, *arnăuțime*, *aromânime*, *artistime*, *avocățime*, *latinime*, *lăutărimime*, *lehime*, *levițime*, *liberălime*, *lipovenime*, *lipscănimime*, *locșitorime*, *locuitorime*, *lotrime* etc.;
- o mulțime de animale: *broștime*, *lăcustime* etc.;
- o mulțime de plante: *nucime*, *lămâime*, *lăptucime*, *stejărimime*, *tufărimime* etc.;
- o mulțime de obiecte: *aurărimime*, *lădițime*, *lănțurime*, *lemnime*, *lesime*, *ruginime*, *scăunime*, *scândurime* etc.

Se observă că în limba română valoarea colectivă a acestui sufix este cea care primează, formându-se cca 200 de substantive colective feminine din totalul de cca 410 derivate cu acest sufix (nu am luat în calcul și variantele lexicale ale acestor derivate). În contextul limbilor romanice, doar româna a păstrat sufixul latin *-imen* în derivate colective; în afara ariei romanice, acest sufix mai apare și în albaneză (Meyer-Lübke 1895: 531).

Cele mai multe derivate cu sufixul colectiv *-ime* au un singur sens cu respectiva valoare, dar există și derivate care au mai multe sensuri dintre care doar unul are un semantism colectiv. De exemplu, *întunecime* 1. „întuneric (adânc), obscuritate”; 2. (despre lună sau soare) eclipsă⁶, 3. (fig.) „lipsă de cultură, barbarie”; 4. (rar) „mulțime nenumărată”; *prostime* 1. „simplitate, modestie; sărăcie”, 2. (învechit) „neștiință, ignoranță, nepricepere” 3. „prostie; ceea ce denotă prostie”, 4. (cu sens colectiv) „oameni de rând, marea masă a populației; s p e c . țărănimime; norod, gloată, mulțime” etc.

Anumite derivate au un sens colectiv dat chiar de cuvântul de bază (vezi *mulțime* < *mult*, *desime* deja menționat etc.), deci sufixul colectiv nu face decât să întărească semantismul colectiv al derivatului.

În ceea ce privește nuanța peiorativă, aceasta poate fi remarcată la unele derivate prin semantismul lor plin [(peior.) *burtăverzime* „burghezie” < *burtăverde* „burghez”, (peior.) *calicime* < *calic*, (peior.) *golănimime* < *golan* etc.] sau în funcție de context [(peior.) *popime* < *popă*].

Din punct de vedere gramatical sufixul colectiv *-ime* se poate atașa: unor substantive (*arăbime* „mulțime de arabi”, *băiețime* „număr mare de băieți”, *țărănimime* etc.), unor adjective (*albăstrime* „oameni de la oraș îmbrăcați în albastru”, *desime* „desiș”, *greime* „mulțime, grosul (oștii)”, *vechime* „oameni din trecut”; unor adverbe (*josime* „oameni de jos”, *călărimime* „oameni călări”); unor verbe (*arzime* „febră”, *însoțime* „grup de oameni, ceată”).

3.2. Sufixul *-iște*

Dacă în cazul derivatelor cu suf. *-ime* caracterul derivat al fiecărui element este ușor identificabil, problema care se pune fiind cea a primitivelor (nu vom discuta acum acest aspect), derivatele cu suf. *-iște* pun problema unei atente disocieri a derivatelor de primitivele care au introdus în limbă respectivul sufix. De exemplu, *gunoiște* este considerat derivat în Sădeanu 1962, MDA, Florescu 2007, câtă vreme în DA și Pascu 1916 este considerat primitiv, împrumutat din vechea slavă. Sau: disocierea omonimelor antrenează niveluri diferite ale raportului dintre

⁶ Precizăm faptul că definițiile au fost simplificate și că spațiul nu ne permite să cităm sursele.

împrumuturile lexicale și derivatul pe teren românesc; de exemplu, *capiște*, cuvânt împrumutat din vechea slavă („templu antic”), *capiște* („căpiță”) împrumutat din sârbă și *capiște* („căpătenie”) derivat pe terenul limbii române.

În funcție de realitatea desemnată, derivatele în *-iște* pot fi grupate în (menționăm cele mai semnificative categorii semantice):

– locul pe care cresc sau sunt cultivate, ori au crescut plante sau arbori: *barabuliște*, *cânepiște* etc.;

– un loc special amenajat sau adăpost natural pentru vite, ori un loc unde își au sălașul animale sălbatice: *bouriște*, *lupiște* etc.;

– un loc având anume caracteristici, calități, trăsături (geomorfologice): *bătețiște*, *goliște* etc.;

– un loc pe care se fac sau pe care s-au făcut anumite construcții, amenajări: *cotiște* „cătun”, *cuptoriște* „loc pentru cuptor” etc.;

– un loc pe care se desfășoară ori unde s-a desfășurat o acțiune, o activitate etc.: *alergăriște* „hipodrom”, *mulgăriște* „loc în care se mulg oile” etc.;

– plantele care cresc, fiind sau nu cultivate, ori fiind depozitate pe un anume teren: *ariniște*, *curpeniște*, *făgiște* etc.;

– un grup de ființe: *roieliște* „mulțime de pui de albină”, *porumbăriște* etc.;

– diverse construcții, amenajări etc. și unele părți ale acestora (cu anumite trăsături caracteristice): *măieriște* „construcție unde se păstrează recolta, uneltele”, *moliște* „porțiune din fagure în care s-au instalat moliile” etc.;

– o acțiune: *măsoariște*, *opreliște*, *pieiște* etc.;

– o stare fizică, sufletească, trăsătură caracteristică etc.: *firiște* „soi, viță”, *liniște* etc.;

– o caracteristică a naturii, fenomen sau stare atmosferică etc.: *noriște* „cer înnorat”, *prigoriște* „caniculă”, *soriște* etc.

Din punct de vedere gramatical, sufixul *-iște* se atașează mai ales unor substantive: *alergăriște* „alergare; hipodrom” < *alergare*; *barabuliște* „ogor pe care se cultivă cartofi” < *barabulă* „cartof”; *făgiște* „pădure de fag” < *fag*; *vraiște* < *vrah* (variantea lui *vraf*) (cu sensul învechit și popular: „snopi de cereale desfăcuți și împrăștiți pe arie pentru a fi treierați cu ajutorul vitelor”). De asemenea, sufixul respectiv se poate atașa și unor adjective – precum: *liniște* < *lin*, *desiște* „teren acoperit cu mulțime deasă de arbori” < *des*, *goliște* „loc neacoperit (de vegetație)” < *gol*, – sau unor verbe: *împărțiște* „împărțire” < *a împărți*, *pribegiște* < *a pribegi*, *zăcăriște* < *a zăcări* „a zăcea”.

4. Valorificare privind istoria limbii române literare: derivatele în *-ime* și *-iște* la Dosoftei.

Evoluția diacronică a cuvintelor derivate cu cele două sufixe și productivitatea acestora marchează semnificativ etapele dezvoltării limbii române literare. Fapt valabil pentru toate articolele DTLR (DA actualizat+DLR), suita de citate cuprinse în corpul fiecărui articol conturează și lexicologic, la nivelul limbii literare, personalități culturale de excepție.

Supunem cercetării câteva aspecte semnificative privind situația derivatelor cu sufixele *-ime* și *-iște* în opera mitropolitului moldovean Dosoftei, considerat ctitorul limbii liturgice românești și al limbajului poetic românesc.

Preocuparea pentru re-valorificarea scrierilor cărturarului [cf. Manea 2006] a scos la iveală, pentru grupul lexical al derivatelor analizate, pe de o parte, valori semantice necunoscute în structura lexicală a cuvintelor deja înregistrate în literatura de specialitate și, pe de altă parte, prime atestări ale derivatelor în discuție, atestări care nu au fost cuprinse în DA și în DLR.

Strict taxonomic, în scrierile cărturarului au fost identificate 20 de cuvinte formate cu sufixul *-ime*: *boierime*, *dăscălime*, *desime*, *frățime*, *grăsime*, *greime*, *grosime*, *iuțime*, *israilime*, *înălțime* (mai ales în forma *nălțime*), *jidovime*, *lărgime*, *lățime*, *lungime*, *mulțime*, *omenime*, *preoțime*, *prostime*, *usturime*, *voinicime* (în forma *vonicime*) și 11 derivate în *-iște*: *alergăriște*, *cinsteliște*, *glodiște*, *goliște*, *gunoiște*, *jărtviște*, *liniște*, *oprealiște*, *priveliște* (în forma *prăviriiște*), *trupiște*, *zăcăriște* (în forma *dzăcăriște*).

Din punct de vedere semantic, și în cazul limbii literare a lui Dosoftei, sufixul *-ime* formează dominant derivate care exprimă: a) calitatea (*desime*, *greime*, *grosime*, *iuțime*, *înălțime*, *lărgime*, *lungime*, *grăsime*) – atașându-se unor adjective, b) o senzație sau o stare – atașându-se unui verb (*usturime*) ori c) are valoare colectivă, desemnând o colectivitate de persoane (*boierime*, *dăscălime*, *frățime*, *israilime*, *jidovime*, *mulțime*, *omenime*, *preoțime*, *prostime*, *voinicime*). Derivate precum *desime*, *mulțime* au sens colectiv dat de însuși cuvântul de bază.

Ca să dăm un singur exemplu, ne vom opri asupra lexemului *greime* (*greu* + suf. *-ime*). Acesta este un cuvânt complex din punct de vedere semantic, care a dispărut din limba literară încă de la începutul secolului al XIX-lea (potrivit DA, acesta apare, ultima dată, în *Fabulele* lui Donici). Atestarea sa regională, dar cu un sens care nu se întâlnește în scrierile vechi, anume „vreme rea, cu furtună”, ne face să presupunem că avea circulație în graiul poporului de la începutul scrisului în limba română. Important este nu numai faptul că termenul este atestat pentru prima oară în opera lui Dosoftei, ci și că scriitorul l-a nuanțat, folosindu-l în construcții pe care astăzi le-am considera figurate: *Atâta-i de dânsul a dzâce, că cu cuvântul pre toț elinii au întrecut. Și osăbi la greimea gândurilor și la lesnea și înflorita graiului* (iunie 46^v/26). Un alt citat: *Apoi Marele Constantin i-au făcut o sântă besearică luminată și scumpă. Și mărirea și greimea aceii besearici scumpe în minunată făptură s-au mutat într-însă sintele lui moștii* (mai 123^v/7) dezvăluie o evoluție semantică a termenului analizat de la sensul „important; decisiv” a adjectivului *greu* (cuvânt *greu*): „importantă, însemnată (deosebită)”.

Cu sufixul *-ime* s-a creat un număr de derivate care s-au păstrat până astăzi în limba literară *boierime*, *dăscălime*, *desime*, *grăsime*, *grosime*, *iuțime*, *înălțime* (mai ales în forma *nălțime*), *jidovime*, *lărgime*, *lățime*, *lungime*, *mulțime*, *preoțime*, *prostime*, *usturime*. Dacă, pe de o parte, cuvinte ca *jidovime*, *omenime*, *voinicime* au avut din limba română veche și până astăzi circulație reală în graiuri, sau un termen bisericesc ca *frățime* este posibil să fie folosit în continuare în cercul clericilor, pe de altă parte, un cuvânt ca *greime* – creație artificială, servind traducerii unor sensuri abstracte ale unor cuvinte slavone sau grecești – a dispărut treptat din limbă din cauza concurenței substantivului *greutate*.

Sufixul *-iște* a fost folosit de mitropolit pentru a crea, de la cuvinte vechi în limbă, o serie de termeni care să-l ajute în traducerea termenilor desemnând unele noțiuni ori unele realități pentru care româna nu avea echivalent. Din acest motiv, multe derivate în *-iște* din limba cărturarului reprezintă formații specifice lui Dosoftei, care nu au avut, probabil, circulație în limbă: *alergăriște, cinsteliște, glodiște, jărtviște, trupiște, zăcăriște*. *Goliște* este un cuvânt care aparține registrului popular, iar *gunoiște, liniște, opreliște, privelyște* rămân termeni bine ancorați în vocabularul limbii române.

Punctual, cercetarea caracterului specific al derivatelor în *-ime* și *-iște* în scrierile reprezentative ale lui Dosoftei, cercetare posibilă datorită materialului adunat, recuperat și prelucrat în proiectul de față, reliefează câteva aspecte.

1. Au fost descoperite prime atestări ale unor cuvinte care sunt înregistrate în DA, DLR din alte texte, mai recente, ale literaturii vechi. Termeni ca: *boierime, grăsimi, grosime* apar în limba scrisă, pentru prima oară, în scrierile lui Dosoftei.

2. S-a găsit un cuvânt nou, *jărtviște*, care nu a fost lucrat în nici unul din dicționarele mari ale limbii române.

3. S-au găsit primele atestări pentru sintagma *liniște veșnică*, ale locuțiunii adverbiale *în prăviriiște* și ale locuțiunii prepoziționale *în prăviriiștea*.

4. S-au descoperit cuvinte care nu au atestări în DA, DLR din opera mitropolitului: *desime, grosime, gunoiște, iuțime, lățime, opreliște*.

5. Au fost recuperate citatele pentru *dăscălime, jidovime*, înregistrate în DA, DLR din antologii de texte românești vechi.

6. Au fost descoperite sensuri ale unor cuvinte care apar pentru prima dată în opera mitropolitului:

– sensul figurat al cuvântului *glodiște*.

– *greime*: „stare, situație etc. dificilă, neplăcută, îngrijorătoare; povară sufletească; dificultate, obligație materială care împovărează situația cuiva”; „cantitate mare”; „forță, profunzime a gândirii; putere de judecată”; „importanță, însemnătate (deosebită)”;

– *grosime*: „diametrul secțiunii transversale a unui corp”;

– *înălțime* (în forma *nălțime*): „grad, rang, demnitate înaltă; persoana care ocupă o asemenea demnitate”; „loc înalt; vârf, culme”;

– *lungime*: „întindere în spațiu a unui drum”; „distanță”; „întindere (prea mare și nejustificată) a unei opere literare, a unui text etc.”;

– *prostime*: „oameni de rând, marea masă a populației; țărănime; norod”.

5. Concluzii

Dacă sufixul *-ime* a avut o productivitate mare în limba literară veche, fiind și în limba contemporană deosebit de activ, mai ales cu valoare colectivă, sufixul *-iște*, extrem de productiv în limba română veche, astăzi și-a încetat, practic, funcționalitatea.

Rezultatele cercetării acestui grant vor fi valorificate într-un volum colectiv, prin care se probează în mod concret eficacitatea acestor tipuri de demersuri lingvistice care valorifică întregul material lexicologic cuprins în DTLR.

Bibliografie

- Carabulea 1959: Elena Carabulea, *-AME și -IME în limba română*, în *Studii și materiale privitoare la formarea cuvintelor în limba română*, vol. I, București, Editura Academiei, p. 65-75.
- Cristea, Răschip et alii, 2007: Dan Cristea, Marius Răschip, Corina Forăscu, Gabriela Haja, Cristina Florescu, Bogdan Aldea, Elena Dănilă, *The Digital Form of the Thesaurus Dictionary of the Romanian Language*, în *Advances in Spoken Language Technology* (edit. Corneliu Burileanu, Horia-Nicolai Teodorescu), București, Editura Academiei Române, p. 195–206.
- DA = *Dicționarul limbii române*, tom I-II, București, Tipografia ziarului „Universul”, Imprimeria Națională, 1913-1937.
- DLR = *Dicționarul limbii române*, Serie nouă, tom VI-XIV, București, Editura Academiei, 1965-2007.
- Dănilă 2007: Elena Dănilă, *Le traitement lexicographique des dérivés aux suffixes collectifs en roumain et en français, en DLR et en TLFi (roum. -ime, -iste et fr. -aille, -erie)*, în *XXV CILPR 2007 Congrès International de Linguistique et de Philologie Romanes. Communications: Résumés*, Innsbruck, Innsbruck University press, p. 572-573.
- Florescu 2006: Cristina Florescu, *Liniște și derivatele pe teren românesc în -iște*, în Volum omagial „Mioara Avram”, București, Editura Academiei, p. 151-160.
- Florescu 2007: Cristina Florescu, *Probleme de semantică a limbii române*, Iași, Editura Universității „Alexandru Ioan Cuza”.
- Haja, Dănilă et alii 2005: Haja, Gabriela, Dănilă, Elena, Forăscu, Corina, Aldea, Bogdan-Mihai, *Dicționarul limbii române (DLR) în format electronic. Studii privind achiziționarea*, Iași, Editura Alfa, publicat și electronic pe www.consilr.info.uaic.ro.
- Iordan 1956: Iorgu Iordan, *Limba română contemporană*, București, Editura Ministerului Învățământului.
- Ivănescu 2000: G. Ivănescu, *Istoria limbii române*, Iași, Editura Junimea.
- Manea 2006: Laura Manea, *Viața și petrecerea svinților. Studiu lingvistic. Partea I*, Iași, Editura Universității „Alexandru Ioan Cuza”.
- Meyer-Lübke 1895: W. Meyer-Lübke, *Grammaire des langues romanes*, II, Paris.
- MDA = *Micul dicționar academic*, București, Univers Enciclopedic, 2001-2003.
- Pascu 1916: Gheorghe Pascu, *Sufixele românești*, București, Edițiunea Academiei Române.
- Sădeanu 1962: Florența Sădeanu, *Sufixele colective din limba română cu specială privire asupra repartiției lor*, în *Studii și materiale privitoare la formarea cuvintelor în limba română*, vol. III, București, Editura Academiei p. 41-88.
- Tușiș, Diaconu 1995: Tușiș, D., Diaconu, L., Barbu, A.M., Diaconu, C., *The Mac-ELU implementation of derivative morphology for Romanian*, Research Report, I.C. I, iunie 1995.

Dérivés dans DTLR. Base lexicale informatisée

Ce travail présente les résultats du projet CNCSIS (code 1609) qui se propose l'acquisition et l'actualisation de l'analyse linguistique d'un group lexical formé des dérivés suffixés en roumain en *-ime* et *-iște*, extraits de DA (complété et actualisé) et de DLR. La base lexicale créée à partir de ces dérivés impose des aspects taxonomiques, lexicologiques, étymologiques et informatiques nouveaux.

La contribution de ce projet au développement de la lexicographie informatisée de la langue roumaine dans l'aire européenne a un caractère édicateur par les implications majeures dans la réalisation du futur eDTLR.

Iași, România