

Veronica PĂCURARU
Institutul de Filologie al AȘM
(Chișinău)

**CONSIDERAȚII PRIVIND
LEXICOGRAFIA INFORMATIZATĂ:
DE LA TRADIȚII LA INOVAȚII
ȘI PERSPECTIVE**

**Considerations of the computerized lexicography:
from traditions to innovations and perspectives**

Abstract. The article tries to validate the thesis that the lexicographical new techniques, which in the description of the language's words relies mainly on the interface and the lexicographical hypertext, will allow undoubtedly the subtilization of the lexicographical treatment of lexical units and the amplification of the spectrum of information relating to their nature and their semiotic, linguistic and functional peculiarities. These new techniques will also assure the elimination of the multitude of annoying inaccuracies and inconsistencies that still persist in the presentation of the lexical units in the printed traditional dictionaries.

Keywords: computerized lexicography, lexicographical corpus, corpora of texts in and for Romanian language, linguistic resources for the Romanian language, printed dictionary, dictionary in electronic version, interface, lexicographical hypertext, semantic beaconing.

După cum se știe, în lexicografia limbilor naturale din întreaga lume, se observă actualmente un amplu și intens proces de schimbare, de modernizare și de optimizare a mijloacelor de elaborare și redactare a lucrărilor lexicografice, în variante tipărită și virtuală, dar și afirmarea unor noi posibilități de accesare și de consultare a acestor surse de informare de către un public interesat tot mai larg. În acest demers inovator, de natură inter- și transdisciplinară, sunt puse în aplicare noile tehnologii informaționale care, după cum s-a demonstrat deja cu prisosință, oferă posibilități incomensurabile pentru crearea de surse lexicografice și resurse lingvistice de factură nouă, favorizând, în special, ameliorarea tratamentului lexicografic al unităților vocabularului limbii, în sensul adecvării lui și al lărgirii spectrului de informații despre unitățile descrise, al precizării statutului semiotic al acestor unități, a semantismului lor, a particularităților de funcționare a lor etc. Această tendință de dezvoltare prezintă totodată noi oportunități de informare (pentru publicul larg) și de cercetare științifică (pentru numeroși cercetători din mediul academic).

Lexicografia română, la fel ca și cea a multor limbi contemporane aparținând unor culturi cu o tradiție multiseclară (în particular, franceză, italiană, engleză, germană etc.), fiind profund marcată de revoluția informațională, se înscrie plenar în acest siaj evolutiv,

iar principalele elemente de schimbare care se pot remarca în domeniu sunt legate mai ales de crearea unor mijloace de informare inovatoare, precum și de punerea în aplicare a unor strategii și instrumente de lucru noi. În acest sens, în ultimii 10-15 ani, au fost desfășurate o serie de proiecte instituționale academice și locale și de inițiative civice de voluntariat, care s-au soldat cu apariția unor surse lexicografice și a unor resurse lingvistice informatizate noi*.

Dacă am lua ca bază criteriile identificatoare ce emerg din caracteristicile pertinente ale noilor izvoare de informare, din însuși modul lor de creare, organizare și funcționare/accesare, aceste surse și resurse electronice noi s-ar lăsa tipizate și clasificate relativ în felul următor:

- corpusuri lexicografice (monolingve și plurilingve);
- corpusuri de texte/discursuri în și pentru limba română;
- resurse lingvistice pentru limba română;
- dicționare informatizate.

Un lucru relevant, care neapărat trebuie subliniat în acest context, este și posibilitatea de utilizare a noilor tehnologii informaționale, a programelor editoriale și a sistemelor de scriere de dicționar create cu ajutorul acestor tehnologii, la elaborarea și redactarea dicționarelor tradiționale, adică a celor editate pe hârtie.

În continuare, vom încerca să prezentăm un tablou, fie și sumar, al acestor surse și resurse electronice pentru limba română, bineînțeles, de pe pozițiile și cu instrumentarul unui lingvist-filolog, și nu al unui informatician.

1. Astfel, dintre tipurile de surse și resurse electronice menționate mai sus, *corpusurile lexicografice monolingve* constituie site-uri web (de obicei cu acces în regim online) care înglobează variantele digitizate ale unor dicționare tipărite de prestigiu ale limbii române, de tipuri diverse (în speță, dicționare explicative, ortografice și morfologice, de neologisme, de antonime și sinonime, etimologice, de termeni speciali pe domenii etc.) și care au fost editate în epoci diferite ale evoluției limbii române actuale. Fiind, în cea mai mare parte, produse pe suport electronic ale unor inițiative private (întrucât sunt create și întreținute de echipe de voluntari), aceste corpusuri lexicografice care sunt aliniat de obicei la nivel de intrare, oferă accesul nelimitat la sursele dezvoltate unui public-țintă larg. Actualmente, pentru limba română există mai multe proiecte importante de acest gen. Firește, în cele ce urmează, voi trece în revistă doar pe cele mai importante dintre ele, care merită a fi reținute grație posibilităților de informare ce le oferă, dar și statisticilor de accesare/consultare a lor.

* Mai multe detalii privind situația în domeniu cu referire la limba română, pot fi găsite în publicațiile unor autori ai proiectelor respective de la Institutul „Al. Philippide” al Ac. Române, filiala Iași, în speță Marius-Radu Clim: *Situația lexicografiei românești în context european*. În: *Philologica Jassyensia*, nr. 2 (16), 2012, p. 259-268; Idem: *La lexicografia rumana informatizada: tendencias, obstáculos y logros*, în vol. *Lexicografia de las lenguas románicas. Aproximaciones a la lexicografía moderna y contrastiva*, vol. II, Editura De Gruyter, 2015, p. 95-110; și Elena Tamba-Dănilă *Despre necesitatea realizării unui corpus lexicografic românesc esențial*. În: *Philologica Jassyensia*, nr. 2 (12), 2010, p. 41-49.

– În această ordine de idei, *Dexonline.ro*, un proiect civic de amploare, creat începând cu 2001 și întreținut inițial de un colectiv de voluntari (cel mai fidel proiectului rămânând a fi astăzi Cătălin Frâncu), s-a impus ca unul dintre cele mai frecventate corpuri lexicografice pe internet. Acest corpus – produs al unor „pasionați de limba română... care au mai multe cunoștințe lingvistice decât omul mediu” (după cum se caracterizează ei înșiși), înglobează 380 000 de definiții, 130 000 de termeni extrași din circa 38 de dicționare tipărite, transpuse pe internet. Este un site cu acces gratuit, având pagina principală (interfața) extrem de simplă (ceea ce favorizează deschiderea ei rapidă) și oferind publicului larg informații variate despre cuvintele vocabularului românesc, care au fost preluate din aceste surse lexicografice importante. Nu este un corpus interactiv; pe *DEX online* nu pune cine vrea ce vrea, toate definițiile sunt verificate și includ sursa, utilizatorul putând alege.

– Un alt site, *Webdex.ro*, reprezintă un subset al *DEX online* și cuprinde un corpus lexicografic cu date preluate atât din *DEXonline*, cât și dintr-o serie de dicționare digitizate ale limbii române, reflectate pe interfața site-ului (în speță: de sinonime, antonime, neologisme, de arhaisme și regionalisme, argou; ortografic, etimologic, academic etc., excluzând dicționarele „de nișă”). Fiind în număr de 27, aceste surse sunt în majoritatea lor apariții lexicografice de dată relativ recentă, care pot fi ușor consultate, întrucât sunt aliniat la nivel de intrare.

– Pe fundalul acestor surse informatizate, se impune în mod special *Corpusul lexicografic românesc esențial (CLRE)*, unul din cele mai ample și polyvalente proiecte instituționale academice de perspectivă, aflat în curs de realizare de către un colectiv de colaboratori ai Institutului de Filologie Română „Alexandru Philippide” al Academiei Române, filiala Iași. Făcând parte din seria demersurilor academice legate de informatizarea limbii române, proiectul dat vizează în principal favorizarea elaborării cu mijloace moderne a *Dicționarului tezaur al limbii române* informatizat. Astfel, conform estimărilor realizatorilor săi, acest corpus de prestigiu va include circa 100 dicționare românești (editate din sec. XVI până în prezent) aliniat la nivel de intrare și la nivel de sens și va avea o interfață și un mediu de programe capabile să permită, pe de o parte, *consultarea interactivă* a acestui corpus, grație construirii unui mediu de programe care să o favorizeze, și, pe de altă parte, *completarea, corectarea și actualizarea* ediției prezente a *DLR*, precum și redactarea viitoarelor ediții ale *Dicționarului Academiei*. Datorită acestor calități ale sale, corpusul în cauză se va constitui într-un cadru modern de cercetare și de activitate lexicografică, ușor adaptabil la o diversitate de obiective.

2. La capitolul *corpusurilor de texte în și pentru limba română* se evidențiază în mod special proiectul de anvergură cu genericul *Corpusul computațional de referință* (sau *corpusul reprezentativ*) *de texte de limba română*, care se constituie într-un program special de cercetare al Academiei Române, ce își propune să realizeze în timp o colecție amplă de date lingvistice, disponibile în format electronic, în forma textelor scrise și/sau a transcrierilor discursurilor vorbite. Publicul-țintă căruia îi este destinată informația cuprinsă în acest corpus de limbă română ține de două categorii majoritare: pe de

o parte – un *public obișnuit*, larg, adică orice cunoscător sau persoană interesată de limba română, din țară sau din străinătate, și, pe de altă parte – un *public specializat*, incluzând profesioniști din domenii diferite, în speță din cel al filologiei și din cel al informaticii, cu preponderență specialiști încadrați în activitățile de prelucrare automatizată a limbajului natural.

Astfel, crearea *Corpusului computațional de referință* urmărește, în mod prioritar, definirea structurii vocabularului românesc și urmărirea acoperirii lui lingvistice, iar în plan strict tehnic, aceasta presupune colectarea de texte de stiluri variate și discursuri transcrise de tipuri diferite, conform unor caracteristici decise; prelucrarea textelor prin tehnologii lingvistice (segmentare, lematizare, dezambiguizare morfolexicală, sintactică etc.); indexarea acestor texte după cât mai multe criterii utile în exploatare; pasarea acestor texte (alias împărțirea lor în unitățile semantico-sintactice sau morfologice componente), pentru extragerea de date lingvistice și de statistici diferite etc., toate aceste operații urmărind să sprijine, în special, cercetătorii la documentare, iar pe lexicografi – la redactarea de dicționare. Căci noile metode de lucru în lingvistica și în lexicografia contemporană accentuează, în particular, necesitatea de construire a teoriei lingvistice, precum și pe cea de elaborare a dicționarelor limbii, pornind de la exemple reperate în texte/discursuri, și nu construite de teoreticieni și de lexicografi. Astfel, această inițiativă de elaborare a unui corpus de texte e menită să-i ajute pe cei aflați în căutarea de exemple-ocurențe ce ilustrează prezența sau absența în vocabularul limbii a unei unități, precum și particularitățile ei de evoluție și de funcționare semantico-sintactică. În acest demers, realizatorii proiectului pentru limba română pot beneficia de experiențele altor limbi contemporane europene, care și-au creat deja pe internet variate corpusuri de texte, cu acces liber sau parțial liber (de exemplu, *Oxford English Corpus*, *British National Corpus* – pentru *engleza britanică*, iar pentru limba franceză în Franța *La base des données textuelles FRANTEXT* și *Les ressources linguistiques informatisées pour l'étude du français*). Comparabilă cu proiectul dat pentru limba română ar fi în acest sens baza digitală de date textuale ale limbii franceze *FRANTEXT*, elaborată inițial pentru a veni în ajutor redactorilor dicționarului *TLF*, ca să constituie actualmente baza textuală cea mai completă a limbii franceze, cuprinzând un număr impunător de circa 4 000 de texte ce aparțin unor stiluri diverse și care datează din secolele XII-XIV până în secolul XXI, și oferind acces de consultare și de hiper-navigare pe bază de abonament.

Momentan, în calitate de soluție inspirată de acest gen, ar putea servi, pentru limba română, informația disponibilă pe platforma *DACOROMANICA*, care include un număr considerabil de resurse textuale, documente ce nu se află sub incidența dreptului de autor și la care accesul este liber, în cea mai mare parte.

În contextul dat, un lucru este cert: într-o epocă în care vizibilitatea și chiar supraviețuirea unei limbi sunt în mare măsură determinate de disponibilitatea acesteia în formă electronică, dezvoltarea unui corpus reprezentativ de date textuale pentru limba română se adeverește a fi un act deosebit de important, atât pentru cercetători și lexicografi, cât și pentru alte categorii de utilizatori interesați.

3. Un loc aparte revine, în ansamblul de izvoare informatizate ale limbii române, categoriei de *resurse lingvistice pentru limba română*, la care se raportează, în particular, site-ul *Archeus.ro*, apărut și întreținut tot printr-o inițiativă privată și înglobând o gamă largă de resurse lingvistice pentru limba română, destinate unui public-țintă eterogen. Astfel, pe lângă facilitățile de căutare a cuvintelor în *DEX online*, site-ul de referință oferă exemple textuale cu ocurențele cuvântului căutat, dar și alte înlesniri pentru limba română, în speță: *corector ortografic și morfologic, program de adăugare a diacriticelor, lematizator* pentru reconstituirea formei-bază a cuvântului etc.

În contextul dat, credem că ar fi cazul ca și pentru limba română să fie creat un tip de resurse lingvistice special, realizat după modelul limbii franceze, care posedă o bază de date speciale, elaborată la *ATILF* – „*La base des mots fantômes*” – a cărei vocație este de a scoate din circuitul lexicografic cuvintele și sensurile „fantomă”, adică cele nedorite în limba franceză, dar și de a exclude lematizările și datările eronate ale unor unități de vocabular înregistrate deja de unele dicționare. Crearea unei atare baze de date ar contribui la sporirea spiritului de ocrotire a geniului limbii române naționale, la protejarea ei contra cuvintelor nedorite, intruse, în vocabularul românesc actual, acel spirit care la francezi e foarte dezvoltat, încă de la începuturile limbii naționale. E suficient să amintim aici grija de cizelare a limbii franceze naționale, care emană chiar din titlul sugestiv „*Défense et illustration de la langue française*” – al manifestului Pleiadei. Or, pentru descifrarea acestui titlu, nu e de ajuns să pui cuvântul *illustration* în relație cu verbul *illustrer* (*a ilustra, a exemplifica*), ci mai bine ar fi ca el să fie asociat cu adjectivul *illustre*, care induce ideea de „*a cizela, a șlefui*”, literalmente, de a face ca limba franceză să fie *ilustră*, căci acesta era adevăratul mesaj și idealul sacru al autorilor manifestului. Pentru limba română, o atare bază de date ar permite a decela creațiile artificiale, efemere, detașându-le net de cele care și-au câștigat dreptul de existență în vocabularul actual și ar constitui o stavilă contra avalanșei de anglicisme și de alte xenisme și barbarisme care invadează actualmente limbajul cotidian.

În loc de concluzie cu referire la resursele limbii române trecute în revistă mai sus, ar trebui subliniat faptul că ele posedă, în totalitate, o caracteristică pertinentă comună: ele reprezintă *variante digitizate, convertite* ale unor lucrări tipărite anterior sau constituie corpusuri electronice de date create tot în baza unor publicații tradiționale, tipărite pe hârtie. Cert e însă că datele furnizate de aceste resurse pot satisface, la moment, necesitățile de informare și documentare ale unui spectru larg de utilizatori, inclusiv cercetători și lexicografi; ele pot fi folosite cu eficiență în procesele de editare a unor dicționare de tip și factură noi – dicționarele informatizate.

Or, viitorul în acest domeniu de informare aparține, fără îndoială, *lexicografiei virtuale*, în speță, *dicționarelor monolingve informatizate*. Edificatoare în acest sens, este și afirmația lingvistului Alain Polguère, unul din membrii de vază ai Centrului *ATILF* (*de Analiză și Tratament Informatic al Limbii Franceze*) care declară: „*La prochaine étape d'évolution de la lexicographie doit être celle d'une lexicographie des dictionnaires virtuels, qui vise en tout premier lieu la construction de structures informationnelles*

à l'image de la structure véritable (ou plausible) des lexiques, structures à partir desquelles peuvent être générées des visualisations aussi bien diagrammatiques que textuelles de l'information lexicale" (în traducere: „Viitoarea etapă de evoluție a lexicografiei trebuie să fie cea a dicționarilor virtuale, care să vizeze, mai întâi de toate, construirea unor structuri informaționale după modelul și imaginea structurii veritabile (sau plauzibile) a diferitor vocabulare, structuri care să permită vizualizarea informației lexicale atât în formă de diagrame, cât și textuală” [1] [traducerea citatului e a noastră, *V.P.*].

4. Și la capitolul *dicționare informatizate monolingve* româna se înscrie cu fermitate în competiția limbilor cu tradiții în domeniu, în special prin proiectul instituțional *Dicționarul limbii române informatizat*, reprezentând varianta virtuală a dicționarului academic eponim și aflat în curs de elaborare la forul academic român, în mare parte după modelul dicționarului limbii franceze *TLFi* (*Trésor de la Langue Française informatisé*). La acest capitol s-ar preta, în parte, și *DEEx.ro*, un proiect mixt, ce pare să îmbine tipurile de *corpus lexicografic* (care conține convertite chiar și unele dicționare de nișă) cu cel de *dicționar electronic*, ambele având avantajul accesării hipertextuale. În această arie se înscrie plenar și proiectul *Dicționarului explicativ de uz general al limbii române (versiune electronică: e-DEGE)*, aflat în desfășurare la Sectorul de lexicologie și lexicografie al Institutului de Filologie al AȘM, care se bazează atât pe o serie de principii lexicografice tradiționale, ce constituie apanajul teoretic și practic al Școlii lexicografice de la Chișinău, fondată de academicianul Silviu Berejan, dar și pe unele principii directe noi, dictate, mai ales, de tendința de subtilizare a prezentării lexicografice a unităților dicționarului, dar și de utilizare la elaborarea lucrării a noilor tehnologii informaționale.

Așadar, în virtutea particularităților de realizare impuse de noile tehnologii informaționale care sunt aplicate la crearea sa, un dicționar informatizat trebuie să aibă ca bază, în primul rând, niște principii teoretice riguroase, care să fie respectate cu strictețe. În acest sens, putem afirma că lingvistul și lexicograful Silviu Berejan a fost un mare vizionar, întrucât, în studiile sale teoretice lexicografice și în preambulul consistent la *DELM-II*, a emis o serie de principii teoretice de structurare și descriere a lexicului limbii naturale, pe care le-a pus în practică la realizarea acestei opere lexicografice, utilizând o metodă nouă, în mare parte revoluționară, de redactare a materialului lingvistic tratat în dicționarul explicativ. Cert este că aceste teze fundamentale ale savantului rămân de mare actualitate mai ales în epoca lexicografiei informatizate, în particular, întrucât, în mare parte, ele se pretează a servi drept principii directorii conceptuale și metodologice în crearea unui dicționar informatizat monolingv.

Pentru conformitate, vom analiza, în paralel, unele caracteristici ale procesului de elaborare și unele principii metodologice (în viziune berejaniană) puse la baza dicționarului monolingv tradițional, în variantă tipărită, și la cea a dicționarului explicativ informatizat.

Cităm astfel din *Principiile...* la *DELM-II*: „[Dicționarul a fost elaborat] pe baza unei **metodici noi de redactare**, ... conform căreia materialul a fost redactat nu după alfabet (adică «de la cuvânt la cuvânt»), ci după clase de cuvinte – morfologice, semantice,

derivative și tematice (adică lucrându-se la fiecare grup de cuvinte, similare din punctul de vedere al tipului de explicație, luat în întregime” [2, p. 9]. Or, la epoca respectivă (în anii '80 ai sec. XX), realizarea acestui deziderat (de altfel, foarte congruent naturii lexico-semantice și specificului morfologic al cuvintelor tratate) s-a adeverit foarte complicată și ea s-a produs în mod artizanal, însă actualmente, odată cu dezvoltarea noilor tehnologii informaționale, în speță a informaticii aplicate, această metodică de tratament al unităților vocabularului se arată unica justă și adecvată, și, grație posibilităților calculatorului care realizează tratamentul informatic al lexicului, ea s-a afirmat ca foarte eficientă și ușor de implementat. Căci anume computerul poate ajuta lexicograful să înlăture cele mai multe dintre dificultățile muncii sale, să sporească rigoarea analizei lexicografice, ce trebuie organizată după un plan foarte bine structurat, menit să țină seama de tipologia materialului lingvistic și de structura și țintele dicționarului informatizat.

În contextul dat, unul din realizatorii dicționarului *TLFi* și un mare teoretician contemporan al lexicografiei monolingve informatizate, Robert Martin, remarcă în 2008, într-un amplu studiu special intitulat „*Perspectives de la lexicographie informatisée*” [3], că prima exigență a unui dicționar informatizat este cea a „unei scriituri lexicografice unificate, căci adesea în lexicografie se întâmplă ca unul și același tip de informație [în ocurență, gramaticală, stilistică, semantică, pragmatică etc. – precizarea ne aparține, *V.P.*] să fie prezentat sub forme diferite... Or, prezentarea unificată nu este nicicum jenantă pentru subiectul uman care consultă dicționarul; din contra, ea are avantajul de a rămâne deschisă pentru o exploatare mecanică” [Idem]. Pentru comparație, în această ordine de idei, Silviu Berejan menționa în 1985, în *Principii... ale dicționarului* evocat mai sus: „Condiția fundamentală, care asigură calitatea unui dicționar, este *unificarea lexicografică* a materialului prezentat. De aceea un dicționar explicativ, alcătuit pe baze cu adevărat științifice, trebuie să prezinte în mod unic sensurile și formele elementelor lexicale descrise” [2, p. 10].

Același R. Martin sublinia, în studiul menționat, că o trăsătură pertinentă a lexicografiei „mecanice” și deci a dicționarului informatizat sunt „formulările explicite”, or, „nimic nu e mai rău în materie de tratament informatizat al lexicului decât o scriitură aluzivă sau ambiguă” [3]. În subsidiar, R. Martin susține: „Exigența explicitării... se va reflecta în lexicografia viitorului nu numai în faptul de a specifica cât mai bine semnificațiile cuvintelor [prin intermediul unor definiții perifrastice de tip aristotelic, – precizarea ne aparține, *V.P.*], dar și în cel de a preciza condițiile care trebuie satisfăcute *contextual* pentru emergența acestor semnificații. Și, în acest sens, e bine venită noțiunea de „*condiții de întrebuițare*”, inițiată de realizatorii *TLFi*” [Idem]. La același subiect, lexicograful Silviu Berejan enunța următorul deziderat: „Cu scopul de a reda cât mai fidel conținutul ontologic al fiecărui cuvânt ca element al limbii, o atenție deosebită s-a acordat formulării sensurilor, ținându-se cont de o serie de prescripțe: – fiecare cuvânt explicat să aibă în mod obligatoriu o definiție analitică, textuală, explicarea numai prin sinonime fiind exclusă ca principiu general; [...] – din fiecare definiție să fie înlăturate elementele ce nu țin de sensul propriu-zis al cuvântului, ci reflectă conținutul îmbinărilor cu acest

cuvânt sau al unui *context mai larg*, deoarece confundarea sensului cuvântului și sensului contextului reprezintă, după cum afirmă lexicologii, unul din cele mai caracteristice neajunsuri ale dicționarilor explicative, existente în prezent” [2, p. 10]. Precizăm aici că această viziune a academicianului Berejan privind importanța reflectării în dicționar a contextului de tipuri variate (context *semantic-sintactic*, care vizează precizarea naturii lexico-semantice a actanților verbelor și adjectivelor); context *pragmatic*, prin care se precizează condițiile de întrebuintare a unităților de vocabular și care este reflectat prin mijloace diferite: fie prin intermediul unor parafraze metalingvistice, descriptive, gen: *se folosește ca formulă de salut la despărțire, mai ales cu copiii; se folosește pentru a exprima neîncredere, îndoială sau surprindere; în Roma antică; în Antichitate și în Evul Mediu; la jocurile de noroc, la fotbal etc.*, fie cu ajutorul unor mențiuni lexicografice de natură pragmalingvistică (*mai ales la pl., la sing., și substantival, și adjectival etc.*) sau de esență pragmastilistică (*fam., verb., afect., poet., iron., peior etc.*), coincide cu teza lui R. Martin privind necesitatea explicitării în dicționarul informatizat a „condițiilor de realizare a sensurilor cuvintelor” [3]. La etapa respectivă, acest precept berejanian a constituit o adevărată inovație în lexicografia explicativă a timpului, iar rigurozitatea cu care ea a fost implementată în lucrarea de referință – una salutară și demnă de urmat.

Evident, caracteristicile sus-menționate ale lexicografiei explicative informatizate vin din *tradițiile* lexicografiei explicative în general, și ele ar trebui să constituie în continuare apanajul atât al dicționarilor tradiționale tipărite, cât și cel al dicționarilor informatizate.

Într-o altă ordine de idei, trebuie constatat însă că un dicționar informatizat se impune și printr-o serie de *inovații*, datorate, în primul rând, aplicării la crearea lui a noilor tehnologii informaționale.

Astfel, lexicografia informatizată transformă dicționarul limbii dintr-o lucrare lexicografică *statică* în una *evolutivă*. Or, se știe că un dicționar tradițional, tipărit, este un obiect imuabil și în mare măsură neterminat, întrucât datele pe care le cuprinde reflectă incomplet starea de facto a limbii la momentul publicării sale, iar, pentru a aduce informația la zi, e necesar fie a publica un supliment al dicționarului respectiv tipărit, fie chiar a edita un nou dicționar (lucru greu de realizat la modul practic, date fiind constrângerile de ordin diferit). Dicționarul informatizat însă, grație tehnicilor noi de creare și construire a sa, permite intervenții ad hoc pentru a actualiza, a re-modela, a corecta, a completa în permanență informația oferită, aducând la zi lista de intrări și datele pe care le prezintă utilizatorilor săi. Un dicționar informatizat este așadar susceptibil de a fixa instantaneu inovațiile din limbă, căci procesul de inovare a vocabularului este unul permanent în fiecare limbă, iar dezvoltarea științifică și explozia informațională contemporană determină transferul rapid de informație și democratizarea limbajului general, ceea ce dictează necesitatea de a schimba viziunea lexicografilor asupra cuvintelor și a termenilor speciali noi: dacă autorii dicționarilor tradiționale ezitau în a da acces acestor elemente de vocabular în registrul unui dicționar de limbă general, atunci dicționarele informatizate, în condițiile în care schimbul de informații și comunicarea se face astăzi cu o viteză

fantastică, trebuie și chiar sunt obligate să adopte o strategie nouă față de unitățile de vocabular noi. Anume în acest sens dicționarul informatizat este evolutiv, ceea ce înseamnă că accesul inovațiilor lexicale în registrul său poate fi permanent și în timp real.

Balizajul (sau *demarcarea*, în termeni obișnuiți) *informațiilor* despre cuvântul descris reprezintă o altă trăsătură fundamentală inovatoare pe care o imprimă noile tehnologii unui dicționar informatizat. Eliberat de forma lineară, *balizajul* ca operație lexicografică primordială se realizează în dimensiuni diverse, comparabile cu „3D”, întrucât beneficiază, pe de o parte, de *interfață*, a cărei structură permite împărțirea informației despre cuvântul descris în zone sau rubrici aparte (cum ar fi: *morfologie*, *definiții sens*, *etimologie*, *relații semantice* (*sinonimie*, *antonimie* etc.), iar, pe de altă parte, de *accesul hipertextual* la datele pe care dicționarul informatizat le oferă utilizatorilor. Aici se impune și sublinierea că *hipertextualitatea lexicografică*, care posedă o arhitectură specifică, oferă lexicografiei informatizate mai multe avantaje capitale. În primul rând, ea favorizează *atomizarea sensului cuvântului* (adică separarea definiției lui, prin intermediul tagurilor (al etichetelor), în termeni constituenți, reprezentând categorii semantice de esență generică și specifică), precum și posibilitatea de consultare a definiției sinonimelor și antonimelor cuvântului definit. În altă perspectivă, *hipertextualitatea* permite și faptul de a face legătura cu alte dicționare informatizate (în special, pentru a compara și completa informațiile referitoare la cuvântul descris), cu bazele de date care au servit ca materie de construcție a ei, cu diverse imagini care vizualizează conținutul semantic al cuvintelor descrise în dicționar etc.

În concluzie, ținem să subliniem în mod special că, în perspectivă lexicografică, noile tehnologii informaționale, care dispun de un spațiu virtual ce pare a fi nelimitat, oferă posibilități incomensurabile pentru îmbunătățirea tratamentului lexicografic al unităților de vocabular, mai ales în sensul precizării statutului și semantismului lor, precum și în al reflectării specificității unităților de limbă, atât a celor noi, recent intrate în limbă, cât și a celor mai vechi. Aceste tehnologii permit (sau chiar favorizează, impun cu certitudine) revederea vechiului principiu lexicentrist de prezentare lexicografică a cuvintelor polisemantice – în cadrul unui singur articol lexicografic, uneori chiar fără mare discernământ – ceea ce duce uneori la includerea eronată în cadrul unui singur articol lexicografic a mai multor unități omonime, care au origini și particularități semantico-funcționale distincte (a se vedea, în acest sens eroarea comisă în cazul cuvântului *magazin*, care în majoritatea dicționarilor explicative tipărite e prezentat drept un cuvânt polisemantic, în timp ce în realitate aici e vorba, fără echivoc, de două omonime etimologice, cu origini distincte: *magasin*, cu sensul de „Local în care se expun și se vând diferite mărfuri” și *magazine* „Publicație periodică (ilustrată) cu material din domenii variate”. E regretabil însă faptul că exemplele de acest gen sunt multiple; lor li se pot adăuga și cazurile deliberat neglijate de omonimie divergentă, conversională, morfologică. Credem, a sosit timpul (și există toate condițiile) ca autorii de dicționare informatizate ale limbii române să ia în considerare principiile de separare a cuvintelor autonome omonime, principii care constituie baza teoriei omonimelor în limba română, susținută și promovată

de echipa de lexicografi formată la Școala lingvistică a academicianului Silviu Berejan. Or, aceasta ar contribui la adecvarea tratamentului lexicografic al unităților lexicale, la aducerea lui în concordanță cu adevăratul statut al unităților funcționale ale limbii.

Nu mai există nicio îndoială că tehnicile informaționale noi, care în descrierea cuvintelor mizează pe interfață și pe structura de hipertext, vor permite subtilizarea tratamentului lexicografic al unităților de vocabular și amplificarea spectrului de informații referitoare la natura și la particularitățile semiotice, semantice și funcționale ale unităților de vocabular descrise în dicționar. Urmează doar ca lexicografii (dar și consumatorii de produse lexicografice moderne, în special informatizate), să învețe a profita cât mai bine și cât mai mult de oportunitățile pe care le oferă noile tehnologii ale informației, iar acestea să fie utilizate în mod judicios și adecvat, uzul lor fiind conjugat cu revizuirea vechilor principii teoretice de interpretare a cuvintelor limbii și cu elaborarea unor principii și metodici noi de tratament lexicografic al unităților de vocabular. Pe această cale, s-ar putea ajunge, mai devreme sau mai târziu, și la înlăturarea multitudinii de imprecizii și inadvertențe supărătoare, care mai persistă încă în prezentarea unităților lexicului limbii în dicționarele tradiționale. Or, aceste imperfecțiuni sunt datorate, în mare parte, multiplelor convenții și constrângeri de ordin diferit (în special, de ordin economico-financiar), pe care sunt nevoite să le accepte autorii de lucrări lexicografice, uneori chiar și din false motive de „economie a spațiului tipografic”. În acest sens, pentru elaborarea unor dicționare performante, care să răspundă unui spectru larg de interese și necesități, se impune un dialog mai strâns între lingviști și specialiștii informaticieni în domeniul procesării limbajului natural, dar și dintre lexicografia științifică (și cea practică) și informatică.

Referințe bibliografice

1. Polguère Alain, *Lexicographie des dictionnaires virtuels*. http://www.ruslang.ru/doc/melchuk_festschrift2012/Polguere.pdf (vizitat 12.11.2015)

2. *Principii de structurare și descriere a lexicului în DELM II*. În: Dicționar explicativ al limbii moldovenești, vol. 2, Chișinău: Redacția principală a ESM, 1985, p. 9-11.

3. Martin Robert, *Perspectives de la lexicographie informatisée*. <http://www.linguistiquefrancaise.org/articles/cmlf/pdf/2008/01/cmlf08332.pdf> (vizitat 17.03.2016).

4. Clim Marius-Radu, *Situația lexicografiei românești în context european*. În: *Philologia Jassyensia*, nr. 2 (16), 2012, p. 259-268.

5. Clim Marius-Radu, *La lexicografía rumana informatizada: tendencias, obstáculos y logros*. În: *Lexicografía de las lenguas románicas. Aproximaciones a la lexicografía moderna y contrastiva*, vol. II, Editura De Gruyter, 2015, p. 95-110.

6. Tamba-Dănilă Elena, *Despre necesitatea realizării unui corpus lexicografic românesc esențial*. În: *Philologia Jassyensia*, nr. 2 (12), 2010, p. 41-49.