Veronica PĂCURARU

Institutul de Filologie al AŞM
(Chişinău)

CONSIDERAȚII PRIVIND LEXICOGRAFIA INFORMATIZATĂ: DE LA TRADIȚII LA INOVAȚII ŞI 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 ştie, în lexicografia limbilor naturale din întreaga lume, se observă actualmente un amplu şi intens proces de schimbare, de modernizare şi de optimizare a mijloacelor de elaborare şi redactare a lucrărilor lexicografice, în variante tipărită şi virtuală, dar şi afirmarea unor noi posibilități de accesare şi de consultare a acestor surse de informare de către un public interesat tot mai larg. În acest demers inovator, de natură inter- şi 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 şi resurse lingvistice de factură nouă, favorizând, în special, ameliorarea tratamentului lexicografic al unităților vocabularului limbii, în sensul adecvării lui şi 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) şi de cercetare ştiințifică (pentru numeroşi cercetători din mediul academic).

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

BDD-A23978 © 2016 Academia de Ştiințe a Moldovei
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 07:50:48 UTC)

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

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

– corpusuri lexicografice (monolingve şi plurilingve);
– corpusuri de texte/discursuri în şi pentru limba română;
– resurse lingvistice pentru limba română;
– dicţionare informatizate.

Un lucru relevant, care neapărat trebuie subliniat în acest context, este şi posibilitatea de utilizare a noilor tehnologii informaţionale, a programelor editoriale şi a sistemelor de scriere de dicţionar create cu ajutorul acestor tehnologii, la elaborarea şi redactarea dicţionarelor tradiţionale, adică a celor editate pe hârtie.

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

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

* Mai multe detalii privind situaţia în domeniu cu referire la limba română, pot fi găsite în publicaţiile unor autori ai proiectelor respective de la Institutul „Al. Philippide" al Ac. Române, filiala Iaşi, în speţă Marius-Radu Clim: *Situaţia lexicografiei româneşti în context european.* În: *Philologica Jassyensia*, nr. 2 (16), 2012, p. 259-268; Idem: *La lexicografía 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; şi Elena Tamba-Dănilă *Despre necesitatea realizării unui corpus lexicografic românesc esenţial*. În: *Philologica Jassyensia*, nr. 2 (12), 2010, p. 41-49.

BDD-A23978 © 2016 Academia de Ştiinţe a Moldovei
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 07:50:48 UTC)

– În această ordine de idei, *Dexonline.ro*, un proiect civic de amploare, creat începând cu 2001 şi întreţinut iniţial de un colectiv de voluntari (cel mai fidel proiectului rămânând a fi astăzi Cătălin Frâncu), s-a impus ca unul dintre cele mai frecventate corpusuri lexicografice pe internet. Acest corpus – produs al unor „pasionaţi de limba română… care au mai multe cunoştinţe lingvistice decât omul mediu" (după cum se caracterizează ei înşişi), înglobează 380 000 de definiţii, 130 000 de termeni extraşi din circa 38 de dicţionare tipărite, transpuse pe internet. Este un site cu acces gratuit, având pagina principală (interfaţa) extrem de simplă (ceea ce favorizează deschiderea ei rapidă) şi oferind publicului larg informaţii 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 definiţiile sunt verificate şi includ sursa, utilizatorul putând alege.

– Un alt site, *Webdex.ro*, reprezintă un subset al *DEX online* şi cuprinde un corpus lexicografic cu date preluate atât din *DEXonline*, cât şi dintr-o serie de dicţionare digitizate ale limbii române, reflectate pe interfaţa site-ului (în speţă: de sinonime, antonime, neologisme, de arhaisme şi regionalisme, argou; ortografic, etimologic, academic etc., excluzând dicţionarele „de nişă"). Fiind în număr de **27**, aceste surse sunt în majoritatea lor apariţii lexicografice de dată relativ recentă, care pot fi uşor consultate, întrucât sunt aliniate la nivel de intrare.

– Pe fundalul acestor surse informatizate, se impune în mod special *Corpusul lexicografic românesc esenţial* (*CLRE*), unul din cele mai ample şi polivalente proiecte instituţionale 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 Iaşi. Făcând parte din seria demersurilor academice legate de informatizarea limbii române, proiectul dat vizează în principal favorizarea elaborării cu mijloace moderne a *Dicţionarului tezaur al limbii române* informatizat. Astfel, conform estimărilor realizatorilor săi, acest corpus de prestigiu va include circa 100 dicţionare româneşti (editate din sec. XVI până în prezent) aliniate la nivel de intrare şi la nivel de sens şi va avea o interfaţă şi un mediu de programe capabile să permită, pe de o parte, *consultarea interactivă* a acestui corpus, graţie construirii unui mediu de programe care să o favorizeze, şi, pe de altă parte, *completarea*, *corectarea şi actualizarea* ediţiei prezente a *DLR*, precum şi redactarea viitoarelor ediţii ale *Dicţionarului Academiei*. Datorită acestor calităţi ale sale, corpusul în cauză se va constitui într-un cadru modern de cercetare şi de activitate lexicografică, uşor adaptabil la o diversitate de obiective.

2. La capitolul *corpusurilor de texte în şi pentru limba română* se evidenţiază în mod special proiectul de anvergură cu genericul *Corpusul computaţional de referinţă* (sau *corpusul reprezentativ*) *de texte de limba română*, care se constituie într-un program special de cercetare al Academiei Române, ce îşi propune să realizeze în timp o colecţie amplă de date lingvistice, disponibile în format electronic, în forma textelor scrise şi/sau a transcrierilor discursurilor vorbite. Publicul-ţintă căruia îi este destinată informaţia cuprinsă în acest corpus de limbă română ţine de două categorii majoritare: pe de

BDD-A23978 © 2016 Academia de Ştiinţe a Moldovei
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 07:50:48 UTC)

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

Astfel, crearea *Corpusului computațional de referință* urmăreşte, în mod prioritar, definirea structurii vocabularului românesc şi urmărirea acoperirii lui lingvistice, iar în plan strict tehnic, aceasta presupune colectarea de texte de stiluri variate şi 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 şi 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 şi în lexicografia contemporană accentuează, în particular, necesitatea de construire a teoriei lingvistice, precum şi pe cea de elaborare a dicționarelor limbii, pornind de la exemple reperate în texte/discursuri, şi nu construite de teoreticieni şi 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 şi particularitățile ei de evoluție şi de funcționare semantico-sintactică. În acest demers, realizatorii proiectului pentru limba româna pot beneficia de experiențele altor limbi contemporane europene, care şi-au creat deja pe internet variate corpusuri de texte, cu acces liber sau parțial liber (de exemplu, *Oxford Enflish Corpus*, *British National Corpus* – pentru *engleza britanică*, iar pentru limba franceză în Franța *La base des données textuelles FRANTEXT* şi *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 şi care datează din secolele XII-XIV până în secolul XXI, şi oferind acces de consultare şi 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 şi la care accesul este liber, în cea mai mare parte.

În contextul dat, un lucru este cert: într-o epocă în care vizibilitatea şi 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 adevereşte a fi un act deosebit de important, atât pentru cercetători şi lexicografi, cât şi pentru alte categorii de utilizatori interesați.

BDD-A23978 © 2016 Academia de Științe a Moldovei
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 07:50:48 UTC)

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 şi întreţinut tot printr-o iniţiativă privată şi înglobând o gamă largă de resurse lingvistice pentru limba română, destinate unui public-ţintă eterogen. Astfel, pe lângă facilităţile de căutare a cuvintelor în *DEX online*, site-ul de referinţă oferă exemple textuale cu ocurenţele cuvântului căutat, dar şi alte înlesniri pentru limba română, în speţă: *corector ortografic* şi *morfologic*, *program de adăugare a diacriticelor*, *lematizator* pentru reconstituirea formei-bază a cuvântului etc.

În contextul dat, credem că ar fi cazul ca şi 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 vocaţie este de a scoate din circuitul lexicografic cuvintele şi sensurile „fantomă", adică cele nedorite în limba franceză, dar şi de a exclude lematizările şi datările eronate ale unor unităţi de vocabular înregistrate deja de unele dicţionare. Crearea unei atare baze de date ar contribui la sporirea spiritului de ocrotire a geniului limbii române naţionale, la protejarea ei contra cuvintelor nedorite, intruse, în vocabularul românesc actual, acel spirit care la francezi e foarte dezvoltat, încă de la începuturile limbii naţionale. E suficient să amintim aici grija de cizelare a limbii franceze naţionale, 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 relaţie 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 şlefui*", literalmente, de a face ca limba franceză să fie *ilustră*, căci acesta era adevăratul mesaj şi idealul sacru al autorilor manifestului. Pentru limba română, o atare bază de date ar permite a decela creaţiile artificiale, efemere, detaşându-le net de cele care şi-au câştigat dreptul de existenţă în vocabularul actual şi ar constitui o stavilă contra avalanşei de anglicisme şi de alte xenisme şi 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 caracteristica pertinentă comună: ele reprezintă *variante digitizate*, *convertite* ale unor lucrări tipărite anterior sau constituie corpusuri electronice de date create tot în baza unor publicaţii tradiţionale, tipărite pe hârtie. Cert e însă că datele furnizate de aceste resurse pot satisface, la moment, necesităţile de informare şi documentare ale unui spectru larg de utilizatori, inclusiv cercetători şi lexicografi; ele pot fi folosite cu eficienţă în procesele de editare a unor dicţionare de tip şi factură noi – dicţionarele informatizate.

Or, viitorul în acest domeniu de informare aparţine, fără îndoială, *lexicografiei virtuale*, în speţă, *dicţionarelor monolingve informatizate*. Edificatoare în acest sens, este şi afirmaţia lingvistului Alain Polguère, unul din membrii de vază ai Centrului *ATILF* (*de Analiză şi 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

BDD-A23978 © 2016 Academia de Ştiinţe a Moldovei
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 07:50:48 UTC)

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 dicționarelor virtuale, care să vizeze, mai întâi de toate, construirea unor structuri informaționale după modelul şi imaginea structurii veritabile (sau plauzibile) a diferitor vocabulare, structuri care să permită vizualizarea informației lexicale atât în formă de diagrame, cât şi textuală" [1] [traducerea citatului e a noastră, *V.P*].

4. Şi 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 şi 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, şi *DExX.ro*, un proiect mixt, ce pare să îmbine tipurile de *corpus lexicografic* (care conține convertite chiar şi unele dicționare de nişă) cu cel de *dicționar electronic*, ambele având avantajul accesării hipertextuale. În această arie se înscrie plenar şi proiectul *Dicționarului explicativ de uz general al limbii române* (*versiune electronică*: *e-DEGE*), aflat în desfăşurare la Sectorul de lexicologie şi lexicografie al Institutului de Filologie al AŞM, care se bazează atât pe o serie de principii lexicografice tradiționale, ce constituie apanajul teoretic şi practic al Şcolii lexicografice de la Chişinău, fondată de academicianul Silviu Berejan, dar şi pe unele principii directoare noi, dictate, mai ales, de tendința de subtilizare a prezentării lexicografice a unităților dicționarului, dar şi de utilizare la elaborarea lucrării a noilor tehnologii informaționale.

Aşadar, 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, nişte principii teoretice riguroase, care să fie respectate cu strictețe. În acest sens, putem afirma că lingvistul şi lexicograful Silviu Berejan a fost un mare vizionar, întrucât, în studiile sale teoretice lexicografice şi în preambulul consistent la *DELM-II*, a emis o serie de principii teoretice de structurare şi descriere a lexicului limbii naturale, pe care le-a pus în practică la realizarea acestei opere lexicografice, utilizând o metodică 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 şi metodologice în crearea unui dicționar informatizat monolingv.

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

Citățm 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,

BDD-A23978 © 2016 Academia de Ştiințe a Moldovei
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 07:50:48 UTC)

derivative şi tematice (adică lucrându-se la fiecare grup de cuvinte, similare din punctul de vedere al tipului de explicaţie, 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 şi specificului morfologic al cuvintelor tratate) s-a adeverit foarte complicată şi ea s-a produs în mod artizanal, însă actualmente, odată cu dezvoltarea noilor tehnologii informaţionale, în speţă a informaticii aplicate, această metodică de tratament al unităţilor vocabularului se arată unica justă şi adecvată, şi, graţie posibilităţilor calculatorului care realizează tratamentul informatic al lexicului, ea s-a afirmat ca foarte eficientă şi uşor de implementat. Căci anume computerul poate ajuta lexicograful să înlăture cele mai multe dintre dificultăţile muncii sale, să sporească rigoarea analizei lexicografice, ce trebuie organizată după un plan foarte bine structurat, menit să ţină seama de tipologia materialului lingvistic şi de structura şi ţintele dicţionarului informatizat.

În contextul dat, unul din realizatorii dicţionarului *TLFi* şi un mare teoretician contemporan al lexicografiei monolingve informatizate, Robert Martin, remarca în 2008, într-un amplu studiu special intitulat „*Perspectives de la lexicographie informatisée*" [3], că prima exigenţă a unui dicţionar informatizat este cea a „unei *scrituri lexicografice unificate*, căci adesea în lexicografie se întâmplă ca unul şi acelaşi tip de informaţie [în ocurenţă, gramaticală, stilistică, semantică, pragmatică etc. – precizarea ne aparţine, *V.P.*] să fie prezentat sub forme diferite… Or, prezentarea unificată nu este nicicum jenantă pentru subiectul uman care consultă dicţionarul; din contra, ea are avantajul de a rămâne deschisă pentru o exploatare mecanică" [Idem]. Pentru comparaţie, în această ordine de idei, Silviu Berejan menţiona în 1985, în *Principii…* ale dicţionarului evocat mai sus: „Condiţia fundamentală, care asigură calitatea unui dicţionar, este *unificarea lexicografică* a materialului prezentat. De aceea un dicţionar explicativ, alcătuit pe baze cu adevărat ştiinţifice, trebuie să prezinte în mod unic sensurile şi formele elementelor lexicale descrise" [2, p. 10].

Acelaşi R. Martin sublinia, în studiul menţionat, că o trăsătură pertinentă a lexicografiei „mecanice" şi deci a dicţionarului 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 susţine: „Exigenţa explicitării… se va reflecta în lexicografia viitorului nu numai în faptul de a specifica cât mai bine semnificaţiile cuvintelor [prin intermediul unor definiţii perifrastice de tip aristotelic, – precizarea ne aparţine, *V.P.*], dar şi în cel de a preciza condiţiile care trebuie satisfăcute *contextual* pentru emergenţa acestor semnificaţii. Şi, în acest sens, e bine venită noţiunea de „*condiţii de întrebuinţare*", iniţiată de realizatorii *TLFi*" [Idem]. La acelaşi subiect, lexicograful Silviu Berejan enunţa următorul deziderat: „Cu scopul de a reda cât mai fidel conţinutul ontologic al fiecărui cuvânt ca element al limbii, o atenţie deosebită s-a acordat formulării sensurilor, ţinându-se cont de o serie de prescripte: – fiecare cuvânt explicat să aibă în mod obligatoriu o definiţie analitică, textuală, explicarea numai prin sinonime fiind exclusă ca principiu general; […] – din fiecare definiţie să fie înlăturate elementele ce nu ţin de sensul propriu-zis al cuvântului, ci reflectă conţinutul îmbinărilor cu acest

BDD-A23978 © 2016 Academia de Ştiinţe a Moldovei
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 07:50:48 UTC)

cuvânt sau al unui *context mai larg*, deoarece confundarea sensului cuvântului şi sensului contextului reprezintă, după cum afirmă lexicologii, unul din cele mai caracteristice neajunsuri ale dicţionarelor explicative, existente în prezent" [2, p. 10]. Precizăm aici că această viziune a academicianului Berejan privind importanţa reflectării în dicţionar a contextului de tipuri variate (context *semantico-sintactic*, care vizează precizarea naturii lexico-semantice a actanţilor verbelor şi adjectivelor); context *pragmatic*, prin care se precizează condiţiile de întrebuinţare a unităţilor de vocabular şi care este reflectat prin mijloace diferite: fie prin intermediul unor parafraze metalingvistice, descriptive, gen: *se foloseşte ca formulă de salut la despărţire*, *mai ales cu copiii; se foloseşte pentru a exprima neîncredere, îndoială sau surprindere*; *în Roma antică; în Antichitate şi în Evul Mediu*; *la jocurile de noroc*, *la fotbal* etc., fie cu ajutorul unor menţiuni lexicografice de natură pragmalingvistică (*mai ales la pl.*, *la sing.*, *şi substantival*, *şi adjectival* etc.) sau de esenţă pragmastilistică (*fam.*, *vorb.*, *afect.*, *poet.*, *iron.*, *peior.* etc.), coincide cu teza lui R. Martin privind necesitatea explicitării în dicţionarul informatizat a „condiţiilor de realizare a sensurilor cuvintelor" [3]. La etapa respectivă, acest precept berejanian a constituit o adevărată inovaţie în lexicografia explicativă a timpului, iar rigurozitatea cu care ea a fost implementată în lucrarea de referinţă – una salutară şi demnă de urmat.

Evident, caracteristicile sus-menţionate ale lexicografiei explicative informatizate vin din *tradiţiile* lexicografiei explicative în general, şi ele ar trebui să constituie în continuare apanajul atât al dicţionarelor tradiţionale tipărite, cât şi cel al dicţionarelor informatizate.

Într-o altă ordine de idei, trebuie constatat însă că un dicţionar informatizat se impune şi printr-o serie de *inovaţii*, datorate, în primul rând, aplicării la crearea lui a noilor tehnologii informaţionale.

Astfel, lexicografia informatizată transformă dicţionarul limbii dintr-o lucrare lexicografică *statică* în una *evolutivă*. Or, se ştie că un dicţionar tradiţional, tipărit, este un obiect imuabil şi 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 informaţia la zi, e necesar fie a publica un supliment al dicţionarului respectiv tipărit, fie chiar a edita un nou dicţionar (lucru greu de realizat la modul practic, date fiind constrângerile de ordin diferit). Dicţionarul informatizat însă, graţie tehnicilor noi de creare şi construire a sa, permite intervenţii ad hoc pentru a actualiza, a re-modela, a corecta, a completa în permanenţă informaţia oferită, aducând la zi lista de intrări şi datele pe care le prezintă utilizatorilor săi. Un dicţionar informatizat este aşadar susceptibil de a fixa instantaneu inovaţiile din limbă, căci procesul de inovare a vocabularului este unul permanent în fiecare limbă, iar dezvoltarea ştiinţifică şi explozia informaţională contemporană determină transferul rapid de informaţie şi democratizarea limbajului general, ceea ce dictează necesitatea de a schimba viziunea lexicografilor asupra cuvintelor şi a termenilor speciali noi: dacă autorii dicţionarelor tradiţionale ezitau în a da acces acestor elemente de vocabular în registrul unui dicţionar de limbă general, atunci dicţionarele informatizate, în condiţiile în care schimbul de informaţii şi comunicarea se face astăzi cu o viteză

BDD-A23978 © 2016 Academia de Ştiinţe a Moldovei
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 07:50:48 UTC)

fantastică, trebuie şi chiar sunt obligate să adopte o strategie nouă faţă de unităţile de vocabular noi. Anume în acest sens dicţionarul informatizat este evolutiv, ceea ce înseamnă că accesul inovaţiilor lexicale în registrul său poate fi permanent şi în timp real.

Balizajul (sau *demarcarea*, în termeni obişnuiţi) *informaţiilor* despre cuvântul descris reprezintă o altă trăsătură fundamentală inovatoare pe care o imprimă noile tehnologii unui dicţionar informatizat. Eliberat de forma lineară, *balizajul* ca operaţie lexicografică primordială se realizează în dimensiuni diverse, comparabile cu „3D", întrucât beneficiază, pe de o parte, de *interfaţă*, a cărei structură permite împărţirea informaţiei despre cuvântul descris în zone sau rubrici aparte (cum ar fi: *morfologie*, *definiţii sens*, *etimologie*, *relaţii semantice* (*sinonimie*, *antonimie* etc.), iar, pe de altă parte, de *accesul hipertextual* la datele pe care dicţionarul informatizat le oferă utilizatorilor. Aici se impune şi 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 definiţiei lui, prin intermediul tagurilor (al etichetelor), în termeni constituenţi, reprezentând categorii semantice de esenţă generică şi specifică), precum şi posibilitatea de consultare a definiţiei sinonimelor şi antonimelor cuvântului definit. În altă perspectivă, *hipertextualitatea* permite şi faptul de a face legătura cu alte dicţionare informatizate (în special, pentru a compara şi completa informaţiile referitoare la cuvântul descris), cu bazele de date care au servit ca materie de construcţie a ei, cu diverse imagini care vizualizează conţinutul semantic al cuvintelor descrise în dicţionar etc.

În concluzie, ţinem să subliniem în mod special că, în perspectivă lexicografică, noile tehnologii informaţionale, care dispun de un spaţiu virtual ce pare a fi nelimitat, oferă posibilităţi incomensurabile pentru îmbunătăţirea tratamentului lexicografic al unităţilor de vocabular, mai ales în sensul precizării statutului şi semantismului lor, precum şi în al reflectării specificităţii unităţilor de limbă, atât a celor noi, recent intrate în limbă, cât şi 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 unităţi omonime, care au origini şi particularităţi semantico-funcţionale distincte (a se vedea, în acest sens eroarea comisă în cazul cuvântului ***magazin***, care în majoritatea dicţionarelor 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 şi se vând diferite mărfuri" şi **magazine** „Publicaţie periodică (ilustrată) cu material din domenii variate". E regretabil însă faptul că exemplele de acest gen sunt multiple; lor li se pot adăuga şi cazurile deliberat neglijate de omonimie divergentă, conversională, morfologică. Credem, a sosit timpul (şi există toate condiţiile) ca autorii de dicţionare informatizate ale limbii române să ia în considerare principiile de separare a cuvintelor autonome omonime, principii care constituie baza teoriei omonimelor în limba română, susţinută şi promovată

BDD-A23978 © 2016 Academia de Ştiinţe a Moldovei
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 07:50:48 UTC)

de echipa de lexicografi formată la Şcoala 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 şi pe structura de hipertext, vor permite subtilizarea tratamentului lexicografic al unităților de vocabular şi amplificarea spectrului de informații referitoare la natura şi la particularitățile semiotice, semantice şi funcționale ale unităților de vocabular descrise în dicționar. Urmează doar ca lexicografii (dar şi consumatorii de produse lexicografice moderne, în special informatizate), să învețe a profita cât mai bine şi cât mai mult de oportunitățile pe care le oferă noile tehnologii ale informației, iar acestea să fie utilizate în mod judicios şi adecvat, uzul lor fiind conjugat cu revizuirea vechilor principii teoretice de interpretare a cuvintelor limbii şi cu elaborarea unor principii şi metodici noi de tratament lexicografic al unităților de vocabular. Pe această cale, s-ar putea ajunge, mai devreme sau mai târziu, şi la înlăturarea multitudinii de imprecizii şi inadvertențe supărătoare, care mai persistă încă în prezentarea unităților lexicului limbii în dicționarele tradiționale. Or, aceste imperfecțiuni sunt datorate, în mare parte, multiplelor convenții şi constrângeri de ordin diferit (în special, de ordin economico-financiar), pe care sunt nevoite să le accepte autorii de lucrări lexicografice, uneori chiar şi din false motive de „economie a spațiului tipografic”. În acest sens, pentru elaborarea unor dicționare performante, care să răspundă unui spectru larg de interese şi necesități, se impune un dialog mai strâns între lingvişti şi specialiştii informaticieni în domeniul procesării limbajului natural, dar şi dintre lexicografia ştiințifică (şi cea practică) şi 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 şi descriere a lexicului în DELM II.* În: Dicționar explicativ al limbii moldoveneşti, vol. 2, Chişinău: 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 româneşti în context european.* În: Philologica Jassyensia, nr. 2 (16), 2012, p. 259-268.

5. Clim Marius-Radu, *La lexicografía rumana informatizada: tendencias, obstáculos y logros*. În: Lexicografia 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: Philologica Jassyensia, nr. 2 (12), 2010, p. 41-49.

BDD-A23978 © 2016 Academia de Ştiințe a Moldovei
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 07:50:48 UTC)