

Monica Busuioc, Alexandru Dan Anghelina, Claudius Teodorescu

INTERFAȚA DE REDACTARE A DICȚIONARULUI LIMBII ROMÂNE

Redactarea în format electronic a *Dicționarului limbii române* (DLR) a presupus, în primul rând, dezvoltarea instrumentelor necesare unui asemenea demers. În acest sens, în cadrul procesului este folosit programul Oxygen, adaptat nevoilor specifice ale lexicografilor. Structurarea informațiilor din DLR este realizată pe bază de standarde internaționale XML (Standard XML), respectiv TEI (Standard TEI).

Avantajele unei astfel de abordări sunt multiple (Busuioc, Teodorescu 2015: 443). Pe de o parte, digitalizarea duce la o structurare riguroasă a informațiilor semantice, stilistice, sintactice și etimologice; pe de altă parte, redactorul este ajutat de diversele automatizări: arborele de sensuri, afișarea listelor de formulări, standardizarea informațiilor legate de etimologie etc.

Asemenea oricărui produs digital, interfața de redactare a suferit pe parcursul anilor modificări menite să îmbunătățească procesul de redactare și structurarea informațiilor.

Dacă, în primele stadii de dezvoltare, interfața de redactare conținea doar formularele ce urmau a fi completate cu date (categorii gramaticale, sensuri, indicații de folosire etc.), în momentul de față, este un *ecosistem* format din formularul de redactare și diverse panouri de administrare a informației, unde se poate interveni în timp real în diferitele liste din interfață (categorii gramaticale, tipuri de indicații de folosire, bibliografie, caractere speciale etc.).

Nu vom insista asupra descrierii globale a secțiunilor prezentate deja în Busuioc, Teodorescu (2015), ci dorim ca, în această parte a lucrării, să prezentăm unele dintre aspectele pe care le considerăm mai interesante pentru redactarea în format electronic a dicționarului – tratamentul variantelor lexicale, al sintagmelor, al locuțiunilor și al expresiilor și etimologia.

Variantele lexicale reprezintă un element important în DLR, pentru că ele „constituie o modalitate de redare a normei lingvistice” (DLR M: XV). În practică, *Dicționarul limbii române* grupează toate variantele unui cuvânt sau sens într-un paragraf separat, la sfârșitul articolului de dicționar:

LR, LXVIII, nr. 2, p. 173–178, București, 2019

MACROBIÓTICĂ s. f.

...

– Și:¹ (învechit) **macrobiótecă** (BARASCH, I. 147), **macroviótică** s. f.

(DLR M).

În acest caz, variantele *macrobiótecă* și *macroviótică* se referă la cuvântul-titlu, însă, în DLR, se pot identifica și situații unde o variantă se referă la unul dintre sensurile intrării:

MÁCRU, -Ă adj., s. f.

1. Adj.

2. S. f.

3. Adj.

...

– Și: (regional, 1) **máclă** adj.

(DLR M).

Nivel: 1

Categorie gramaticală:

Parte de vorbire*: adj. Tip: Număr:

Slab, fără grăsime și fără oase.

Indicații privind folosirea: despre carne, în special despre carnea de m:

Atestare: Autor

Citat:

Izvor: unknown Căutare Dată:

Localizare în sursă:

Variantă lexicală

Izvor: ALR I (DA) ALR L... | ALR I Căutare Dată: 739/308

Localizare în sursă:

Categorie gramaticală:

Parte de vorbire*: adj. Tip: Număr:

Fig. 1. Secțiunea *sens* pentru cuvântul-titlu.

¹ În noua ediție a dicționarului, formularea „– Și:” va fi înlocuită cu „– Var:”, pe care o considerăm mai transparentă.

Decizia de a redacta variantele corelat cu elementele de care sunt legate conceptual (cuvânt-titlu sau sens) pare, la prima vedere, neconvențională, fiind însă motivată după cum urmează.

Organizarea interfeței de lucru a presupus identificarea celor mai potrivite soluții pentru a exista un echilibru între forma finală, tipărită, și o logică a redactării. Chiar dacă toate variantele sunt grupate în paragraful destinat acestora, pentru redactare s-a decis ca varianta specifică unui anumit sens să fie inserată direct în subsecțiunea dedicată sensului respectiv. Astfel, corelarea sens-variantă se face direct, făcându-se în același timp și distincția între variantă a cuvântului-titlu și variantă a unui sens.

Un alt aspect pe care dorim să îl menționăm este tratamentul sintagmelor, al locuțiunilor și al expresiilor. Identificarea exactă a fiecărei sintagme sau expresii a presupus adaptarea normelor de redactare la noile cerințe de digitalizare. Standardele adoptate în redactarea în format electronic a dicționarului nu reușesc să ofere soluții pentru toate situațiile din DLR, fiind uneori necesare abordări alternative, care, în opinia noastră, nu modifică esența dicționarului.

În primul rând, în practica lexicografică existentă în DLR, atunci când cuvântul-titlu este parte componentă a unei sintagme, fără ca sensul acestuia să sufere modificări, sintagma este menționată în paranteza destinată indicațiilor, după cum se poate vedea la **DÁFIN**:

DÁFIN s. m. (Bot.)

1.

2. (Olt. și prin Munt.: și în sintagma *dafin creț*, LEXIC REG. 41) Salcâm (1) (*Robinia pseudacacia*).

(DLR D)

Imposibilitatea de a izola sintagma în cadrul parantezei a necesitat găsirea unei soluții care să păstreze legătura semantică dintre cuvântul-titlu și sintagma din care face parte. În acest sens, s-a ales, ca soluție informatică, repoziționarea și marcarea sintagmei sub sensul respectiv:

DÁFIN s. m. (Bot.)

1.

2. (Olt. și prin Munt.) Salcâm (1) (*Robinia pseudacacia*).

◆ Sint². **Dafin creț** v. **dafin** (2).

1960 *Lexic Reg.* I

În al doilea rând, situațiile în care o structură are o formă de tipul ◆ Expr. **A da** (sau) **a face** (ceva) **pe daibojeală**, suntem de părere că reprezintă, de fapt, două

² În noua ediție a *Dicționarului*, sintagmele vor fi izolate cu indicația **Sint**.

expresii **A da pe daibojeală**, respectiv **A face pe daibojeală** comasate într-o singură formulare, pe baza aceluiași sens. Din acest motiv, a fost luată decizia segmentării structurii inițiale în două unități distincte, care pot avea atașate diverse informații, precum indicațiile de folosire (rar, înv. etc).

The image shows a software interface for semantic expressions. At the top, there are several tabs: 'Cat. gram.', 'Ind. fol. globală', 'Trimitere', 'Variantă lexicală', and 'Etimologie'. Below these, there are two identical-looking sections. Each section has a title 'Unitate semantică subsumată:' followed by a red minus sign and a green plus sign. Underneath each title are four sub-tabs: 'Ind. fol.', 'Tip verbal', 'Abreviere', and 'Etimologie'. Below the sub-tabs is a 'Tip:' section with several checkboxes: 'proverb', 'zicătoare', 'ghicitoare', 'sintagmă', 'expresie', 'compus', 'loc. adj.', 'loc. adv.', 'loc. prep.', 'loc. conj.', and 'loc. vb.'. The 'expresie' checkbox is checked in both sections. Below the 'Tip:' section is a 'Conținut:' field. The first section's content is 'A da pe daibojeală' and the second's is 'A face pe daibojeală'.

Fig. 2. Secțiunea *expresie semantică subsumată*.

Așadar, vechea metodă de redactare a fost adaptată pentru ca fiecare sintagmă, locuțiune sau expresie să poată fi izolată și marcată ca atare.

Din punctul de vedere al redării, deci al utilizatorului dicționarului, expresiile de mai sus vor fi izolate printr-o bară verticală și afișate astfel: **◆ Expr. A da pe daibojeală | A face pe daibojeală** = a da sau a face ceva pe datorie.

Un ultim element pe care îl prezentăm în acest articol este legat de redactarea etimologiilor.

S-a procedat la îmbunătățirea definirii etimologiilor, prin împărțirea lor în patru tipuri:

1. etimologia cuvântului-titlu;
2. etimologia unei variante lexicale;
3. etimologia unui sens;
4. etimologia unei sintagme, expresii etc. (în cazul unor calcuri).

Astfel, în cadrul redactării, tratarea etimologiilor se face în interiorul subsecțiunii de care aparțin, așa cum este cazul și pentru variantele lexicale.

Etimologia unui sens sau cea a unei sintagme sau expresii nu este indicată la sfârșitul articolului de dicționar, ci imediat lângă acestea:

TĂBLIȚĂ s. f.

...

II.

...

3. (Învechit, la pl.; calc după fr. *tablette*) Carnetele în care se notează diverse fapte, maxime etc. (DLR T)

◆ Expr. (Rar; calc după fr. *Pierre de scandale*) **Piatră de scandal** = motiv de nemulțumire, de conflict.

(DLR P)

Având în vedere aceste lucruri, interfața a fost dezvoltată în așa fel, încât redactorul să poată avea în același loc *casetele* necesare, evitând astfel navigarea suplimentară către aceste elemente.

De asemenea, s-a avut în vedere caracterul facultativ al anumitor componente. Dacă etimologia cuvântului-titlu este obligatorie (prezentă în finalul primului ecran al interfeței), etimologia sensurilor se impune numai în anumite situații. Din acest motiv, s-a ales implementarea unui sistem de butoane care să activeze, după specificul fiecărei redactări, componentele respective.



Fig. 3. Secțiunea *etimologie* pentru cuvântul-titlu.

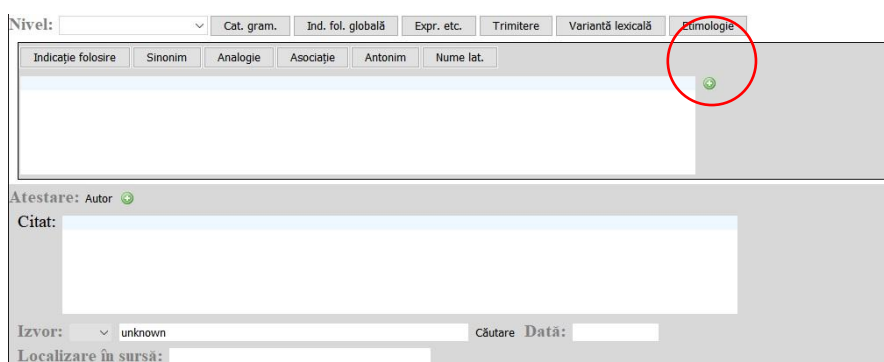


Fig. 4. Secțiunea *etimologie* pentru sens.

În opinia noastră, redactarea *Dicționarului limbii române* în format electronic este benefică pentru potențialii utilizatori ai dicționarului, prin caracterul structurat al informației, fiind posibile cercetări variate (semantice, etimologice, stilistice).

Un dicționar redactat în format digital, pe baza unor standarde internaționale, va contribui la vizibilitatea lexicografiei românești în lume.

BIBLIOGRAFIE

Busuioc, Monica, Claudiu Teodorescu, 2015, „Informatizarea *Dicționarului limbii române* – un vis pe cale de realizare”, în Marius Sala, Maria Stanciu Istrate, Nicoleta Petuhov (ed.), *Lucrările celui de-al cincilea Simpozion Internațional de Lingvistică, București, 27–28 septembrie 2013*, București, Editura Univers Enciclopedic Gold, p. 441–445.

SIGLE

DLR D: *Dicționarul limbii române. (DLR)*. Serie nouă. Tomul I. Partea a 3-a. Litera D. D – Deînmulțit: 2006.

DLR M: *Dicționarul limbii române. (DLR)*. Serie nouă. Tomul VI. Litera M: 1965–1968.

DLR P: *Dicționarul limbii române. (DLR)*. Serie nouă. Tomul VIII. Litera P: Partea a 2-a: *Pe–Pînar*, 1974.

DLR T: *Dicționarul limbii române. (DLR)*. Serie nouă. Tomul XI. Partea a 2-a: Litera T. *T–Tocăliță*: 1982.

Standard TEI: www.tei-c.org (versiunea 3.6.0; accesat la data de 17.07.2019)

Standard XML: www.clarin.eu/content/standards-and-formats (versiunea 1.0; accesat la data de 17.07.2019)

THE DRAFTING INTERFACE OF THE ROMANIAN LANGUAGE DICTIONARY

(Abstract)

The paper aims to broadly describe the drafting interface of the *Romanian Language Dictionary* (DLR) and to provide some insights regarding the process of compiling such a dictionary.

Cuvinte-cheie: dicționar electronic, Oxygen, etichete, TEI.

Keywords: digital dictionary, Oxygen, tags, TEI.

Institutul de Lingvistică al Academiei Române
„Iorgu Iordan – Alexandru Rosetti”, București
monica.busuioc@yahoo.com
anghelinaalexandru@yahoo.com
Universitatea din Heidelberg, Germania
claudiu.teodorescu@gmail.com