

## CLASELE DE MARCHERI DE DEFINIȚII PENTRU ANALIZA SEMANTICĂ A INTRĂRILOR DLRLC

BURCA Eugenia

Universitatea Agrară de Stat din Moldova,

Institutul de Matematică și Informatică al AȘM

Recenzent: BOIAN E. , dr.

Cuvinte cheie: *marcheri de sens și de definiții, resurse lingvistice, parsare, intrare de dicționar, arbore de sensuri.*

În lucrarea de față a fost utilizată metoda de parsare bazată pe configurații SCD (Segmentare-Coeziune-Dependență) [1] pentru transformarea eficientă a intrărilor (ce conțin text neprelucrat) din Dicționarul Limbii Ruse Literare Contemporane (DLRLC) în 20 volume [2] într-o resursă lingvistică indexabilă. Strategia care stă la baza acestei metode reunește două configurații diferite de parsare: una, care identifică și extrage pentru fiecare intrare de dicționar arborele de sensuri principale și secundare și altă configurație, care parsează fiecare nod din arborele de sensuri cu scopul de a clasifica *definițiile* din dicționar, ce descriu sensuri principale și secundare.

Un *marcher* reprezintă o graniță pentru o categorie lingvistică specifică. Recunoașterea marcherilor și construirea arborelui de sensuri a unei intrări din DLRLC a fost realizată în lucrarea [3]. Clasele de marcheri folosite pentru identificarea arborelui de sensuri pentru sensuri principale sînt: *cifrele romane (I, II, III,...etc.)*, *cifrele arabe urmate de punct (1., 2., 3.,... etc.)*, *două bare oblice (//)*. Pentru sensuri secundare se utilizează: *rombul gol („◇”)*, *traversa („=”)*, *tilda („~”)*, *enumerarea literală - de tip litere chirilice, după care urmează paranteza rotundă ( a), б), в),...etc.)*.

În urma analizei conținutului semantic al sensurilor pentru intrările din DLRLC s-a depistat următorul set specific de clase de marcheri de definiție: *MorfDefs*-definiții morfologice; *RegDefs*- definiții scrise cu font regular; *BoldDefs*- definiții scrise cu bold; *ItalDefs*- definiții scrise cu italic; *SpecDefs*- definiții ce conțin specificații; *SpSpecDefs*- definiții scrise cu litere spațiate, ce conțin anumite specificații; *DefExems*- exemple la definiții, cu scopul de a întregi sensul unei definiții. Acești marcheri de definiție coincid întru-totul cu marcherii definițiilor propuși pentru Dicționarul Tezaur al Limbii Române[1].

### Bibliografie:

- 1 N. Curteanu, A. Moruz, D. Trandăbăț, C. Bolea, M. Spătaru, M. Husarciuc. Parsarea arborilor de sensuri și segmentarea la definiții în dicționarul tezaur eDTLR. – în *Lucrările atelierului Resurse lingvistice și instrumente pentru prelucrarea limbii române*. Iași, 19-21 noiembrie 2008, Editura Universității „Alexandru Ioan Cuza”, Iași, pp. 65-74.
2. Словарь современного русского литературного языка. В 20 томах. Издательство: М.: Русский язык; Издание 2-е, перераб. и доп. Переплет: твердый; 864 страниц; 1991 - 1994 г. ISBN: 5-200-01068-3.
3. E. Burca. Achiziționarea cunoștințelor lingvistice prin parsarea arborilor de sensuri din intrările de dicționar. – în *lucrările conferinței a 3-a internațională “TELECOMMUNICATIONS, ELECTRONICS AND INFORMATICS” ICTEI 2010*, ISBN 978-9975-45-082-9, ISBN 978-9975-45-136-2(VOL. II), CHIȘINĂU, MOLDOVA, 20-23 MAI, 2010.