

TERMINOLOGIA HRANEI ÎN TEXTE ROMÂNEȘTI ALE NOULUI TESTAMENT DIN SECOLELE AL XVI-LEA ȘI AL XVII-LEA *

CĂTĂLINA MĂRĂNDUC^{1,2}, MIHAELA MOLDOVEANU¹, CENEL AUGUSTO PEREZ²

¹Institutul de Lingvistică „Iorgu Iordan – Al. Rosetti”

²Facultatea de Informatică, Universitatea „Alexandru Ioan Cuza” din Iași
catalinamaranduc@gmail.com, mmoldoveanu82@yahoo.com,
perez_cenel_augusto@yahoo.com

Abstract: In this paper, we talk about the RoDia project, *i.e.* the creation of a corpus type treebank in Dependency Grammar, to illustrate the old Romanian language (from the 16th to the 19th century). We have annotated the Gospels of the *New Testament* from Alba Iulia (1648), the first edition printed in Romanian. We have the intention to annotate the *Acts of Apostles*, the second part of the book and to affiliate this corpus at international projects. At the same time, we aim to give examples of research that can be carried out using this corpus. There have been described food terminology, meanings and symbols, based on quotations extracted from the similar books printed by Coresi, from the 1688 *Bible*, and from the 1648 *New Testament*.

Keywords: treebank, old Romanian, food terminology, figurative meaning, advanced searches.

1. Introducere

Corpusul de tip *treebank*, cu adnotare morfologică și sintactică automată corectată manual, este un proiect demarat la Facultatea de Informatică a Universității „Alexandru Ioan Cuza”, în 2007. La început, acesta conținea traduceri din limba engleză și citate analizate în culegeri cu scopuri didactice, deci limbă română contemporană standard. Ulterior, s-au adăugat tot mai multe citate din *Dicționarul limbii române* (DLR) sau din bibliografia lui (din toate epocile și în toate stilurile) și, de asemenea, din comunicarea social-media, folclor regional, texte vechi, devenind un corpus balansat, adică încercând să cuprindă toate stilurile funcționale ale limbii române.

În general, în Procesarea Limbajului Natural (NLP), *corpora* au rolul de a asigura antrenarea instrumentelor de procesare automată. Totuși, cât din comunicarea într-o limbă este standard sau respectă normele de corectitudine? Nici chatul, nici limbajul familiar, popular sau regional, nici literatura, nici jurnalismul nu sunt obligate să o

* *Food Terminology in Romanian New Testament Texts from the 16th and 17th Centuries.*

facă și își permit licențe poetice și inovații. Atunci, dacă vom antrena instrumentele cu corpus standardizat, cum vor procesa ele un limbaj natural non-standard? Pe de altă parte, lingviștii sunt tot mai mult interesați de realizarea unor studii asupra limbii vechi cu mijloace de analiză moderne, iar informaticienii sunt interesați de digitizarea moștenirii culturale, care presupune nu doar scanarea cărților vechi, ci și citirea automată a informației din ele.

Venind în întâmpinarea acestor necesități, am decis să creăm un sub-corpus consistent al *trebankului* Facultății de Informatică, cuprinzând limbă română din texte scrise între secolele al XVI-lea și al XIX-lea, și pe baza acestui corpus, pe care l-am numit RoDia (Română Diacronică), să antrenăm instrumente de procesare care să poată digitiza limba română veche.

În prezent, acest corpus de limbă veche cuprinde cinci documente XML care conțin, în 5028 de arbori sintactici corecți manual, textul integral al celor patru evanghelii din *Noul Testament* de la Alba Iulia (1648), precum și un document cu 695 de citate din perioada cuprinsă între secolele al XVI-lea și al XVII-lea. Alte documente – *Legile lui Caragea* (1818), *Cronica lui Ion Neculce* (1743), *Floarea Darurilor* (1595) – sunt în curs de corectare manuală. Întregul corpus RoDia are o variantă adnotată semantic și una cu cuvintele din text în caractere chirilice editabile, obținute cu ajutorul unui program de citire a literelor chirilice, creat la Chișinău (pentru aceasta, vezi Cojocar/ Colesnicov/ Malahova 2017).

Programul a început să funcționeze în august 2016 și este o premieră absolută. Este vorba despre o carte tipărită în secolul al XVII-lea, scrisă cu caractere chirilice românești, citite de un OCR (Optical Character Recognizer) și transformate în text editabil (*Fig. 1*). Textul recunoscut are o acuratețe de cca 65% și trebuie corectat manual. Acuratețea acestuia va crește pe măsură ce dicționarul programului va conține mai multe cuvinte scrise cu caractere chirilice, care provin din corectarea manuală.

În a doua etapă, programul OCR este completat cu un convertor de litere chirilice în litere latine. Caracterele folosite, create tot la Chișinău, sunt internaționale și se pot folosi și în formatul XML. Varianta cu caractere chirilice a corpusului RoDia este construită pentru a se putea introduce informație în dicționarul programului OCR, dar și pentru a antrena, în viitor, instrumente de procesare a limbii române vechi scrise cu caractere chirilice.

În noiembrie 2016 am creat o clonă a analizorului morfologic automat (folosit pentru texte contemporane standard) al Facultății de Informatică. Aceasta se numește OldRoPOSTagger (Analizor al părților de vorbire pentru limba română veche) și are o acuratețe de 91,66%. Un program auxiliar introduce permanent în lexiconul lui cuvintele nou întâlnite, după ce analiza lor automată este corectată manual; astfel, acuratețea acestuia va crește odată cu corpusul (Mărănduc/ Simionescu/ Cristea 2017).

În plan internațional, aceste două programe pentru limbile vechi s-au dovedit a fi la același nivel cu cele create în alte țări (OCR pentru litere gotice, grecești, copte,

vechi slave, analizoare morfologice pentru texte din secolul al XV-lea sau pentru limbi antice)¹.

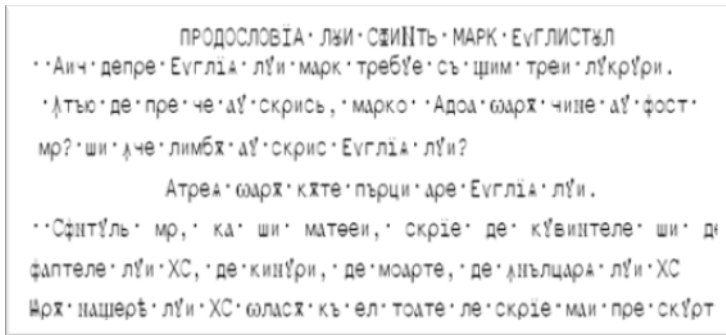
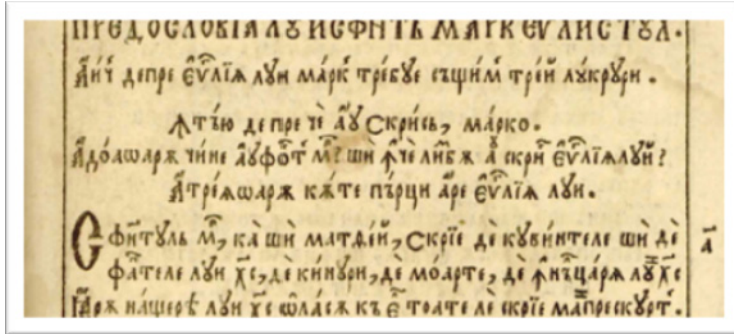


Fig. 1. Fragment din *Predoslovie* la *Evangelia după Marcu* și ocerizarea lui.

2. Prezentarea textelor

În cele ce urmează, ne propunem să facem cunoscut corpusul menționat și să oferim un exemplu de utilizare a acestuia în scopul unor cercetări lingvistice. În acest sens, am întreprins o comparație a vocabularului din corpusul nostru cu cel din textele lui Coresi și cu cel din *Biblia de la 1688*, alegând terminologia hranei, purtătoare de semnificații multiple. Urmărim să vedem dacă în trei traduceri de origine diferită ale textelor sacre în limba română sunt folosiți aceiași termeni din sfera semantică amintită și dacă ei sunt purtători ai acelorași semnificații.

¹ The 18th International Conference on Computational Linguistics and Intelligent Text Processing (Budapest, 17-23 aprilie 2017) și Digital Access to Textual Cultural Heritage International Conference (Göttingen, 1-2 iunie 2017).

Pentru a obține texte paralele, am reunit două cărți ale diaconului Coresi, și anume *Tetraevanghelul* (1561) și *Lucrul Apostolesc* (1563), care conțin fragmente din Noul Testament; din *Biblia de la 1688* am selectat numai textele noutestamentare². În continuare, prezentăm pe scurt textele selectate:

a) *Tetraevanghelul* (= T) este a doua carte românească tipărită de Coresi (1560-1561);

b) *Lucrul Apostolesc* (= LA), tipărit în 1563 de Coresi, constituie, împreună cu *Tetraevanghelul*, o mare parte din *Noul Testament*;

c) *Noul Testament* de la Bălgrad (= NT) este prima traducere integrală în limba română a Noului Testament și a fost tipărit în 1648, în timpul principelui Gheorghe Racoczi I, sub coordonarea mitropolitului Simion Ștefan; traducerea s-a făcut după *Vulgata* publicată în 1587, la Venetia;

d) *Biblia de la 1688* sau *Biblia lui Șerban Cantacuzino* (= B) este prima traducere completă a Bibliei în limba română. Traducerea s-a bazat pe versiunea lui Nicolae Milescu a *Septuagintei*, revăzută de frații Radu și Șerban Greceanu.

Este vorba, așadar, despre traduceri diferite, după surse diverse. Influențele ideologice suferite de traducători sunt și ele divergente. Textele lui Coresi prezintă influențe luterane, iar *Noul Testament* de la 1648, influențe calvine. La nivel lexical, toți traducătorii se străduiesc să evite regionalismele și se pare că provin din zone geografice ale țării unde acestea se manifestă foarte puțin (București, Brașov, Târgoviște, Țara Românească). Toate textele prezintă influențe slave, întrucât slavona fusese anterior limbă de cult și traducătorii se străduiau să stabilească corespondențe slavo-române necesare preoților.

Textele au fost citite simultan cu ajutorul unui program numit LUCON, care are acces liber pe pagina SURCEFORGE³, cu care se pot efectua simultan căutări în toate documentele deschise (pot fi sute de cărți în format text, incluse într-un folder creat special).

3. Terminologia hranei. Analiză comparativă pe baza unui corpus electronic adnotat sintactic și semantic

3.1. ALUAT – subst. moștenit din lat. *allevantum* ‘ridicat, înălțat’ (DLR, *s.v.*). Programul LUCON ne-a oferit o statistică a ocurențelor, în urma analizei căreia putem preciza că acest termen apare în textele-sursă în mai multe forme, cu următorul număr de ocurențe pentru fiecare formă: *aluat* (LA 6, NT 5, B 6), *aluatul* (LA 6, NT 12, B 6), *aluatului* (NT 2, B 2, T 1); termenul este folosit cu sens figurat (prin propovăduire

² Textele selectate sunt printre primele traduceri românești ale textelor sacre și sunt editate de filologi. Acestea sunt stocate într-un corpus deținut de Institutul de Lingvistică „Iorgu Iordan – Al. Rosetti” din București.

³ Vezi <https://sourceforge.net/directory/os:windows/?q=Lucon> (17.08.2017). Programul este creat de informaticianul Cătălin Mititelu.

influența unei religii crește și se răspândește”), este utilizat însă și cu sens negativ în NT.

NT: Asămănată iaste împărăția cereiurelor *aluatului* carele ia muiarea de-l acopere în făină de trei măsuri, până să dospeaște tot (Mat. 13:33);

T: Cui voi asemana împărăția lu Dumnezeu? Aseamnă-se *aluatului* ce ia muiarea, acoapere în făinina de trei măsuri până se năcreaște tot (Mat. 13:33);

NT: Căutați și vă păziți de aluatul fariseilor și a saducheilor (Mat. 16:5); Atuncea înțeleseră că nu le-au zis să să ferească de *aluatul* pîinei, ce de învățăturile fariseilor și a saducheilor (Mat.16:12).

Creșterea aluatului este, deci, un simbol al propagării influenței, indiferent dacă aceasta este pozitivă sau negativă. Se observă că, inițial, ucenicii confundă acest înțeles parabolic cu simbolistica pâinii.

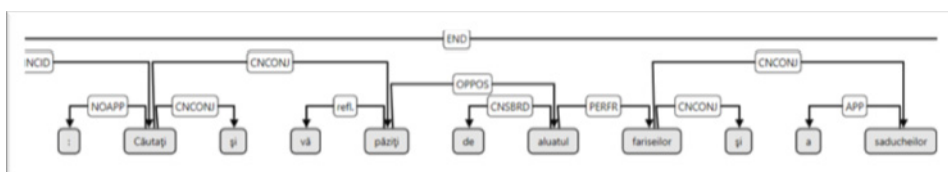


Fig. 2. Mat. 16:5 în adnotare semantică.

3.2. BĂUTURĂ – subst. moștenit din lat. **bibitura* (DLR, *s.v.*). *Bentură* este o variantă regională din Ardeal; în NT are 7 atestări; termenul standard *băutură* are 3 atestări în NT și 8 în B. Unele ocurențele denotă un sens spiritual, de exemplu: „Și toți aceaș mîncare duhovncească au mîncat, și toți aceaș *băutură* duhovncească au băut” (B, 1Cor. 10:4). Alteori, termenul este folosit cu sensul propriu, *i.e.* ‘parte din cele necesare trupului’: „Și cine va da *băutură* unuia den cești mai mici numai un pahar de apă reace, în nume de ucenicu, adevăr grăiesc voao, nu va piarde plata lui” (B, Mat. 10:42). Pluralul *băuturi* apare în LA 1, NT 2 și B 2. Pare să predomine aici sensul negativ, *i.e.* ‘element al unei vieți petrecute în păcate și desfătări trupești’: „Ne-am petrecut anii vieții voia limbilor făcînd [...] întru pohtiri, întru beție, întru cimpoaielor glasure, întru *băuturi*, întru necurăție, slujbele idolilor” (LA, Tit 9:10).

3.3. BRÂNZĂ – subst. provenit de la numele orașului Brienz; cuvântul desemna, la început, un anumit tip de brînză (DLR, *s.v.*); această explicație s-ar potrivi unui neologism. Brîncuș (1983, p. 121) și DELR, *s.v.* îl consideră element de substrat. Este atestat prima dată într-un document din 1356-1381, menționat în CUV. B, vol. I, p. 190. În T (1561) îl găsim în *Pravila svenților oteți după învățatura a marelui Vasilie*: „Cine se spurcă în postul mare cu carne sau cu *brîndză*, elu să postească den Dumineca Tomeci pînă în Rusali” (T, 209^r). Cuvântul are aici sens propriu, *i.e.* ‘hrană care nu respectă canoanele postului’.

3.4. BUCATE – subst. moștenit din lat. *buccata* (DLR, *s.v.*). Apare în corpusul nostru cu următoarele atestări: *bucate* (LA 12, NT 15, B 8, T 14), *bucatele* (LA 7, NT 3, B 2,

T 1), *bucatelor* (LA 3, NT 2, B 1). Sensul obișnuit este de ‘hrană necesară trupului, de care nu ar trebui să ne îngrijim’. La singular are sens similar celui din limba contemporană: „Răspunse Iisus: «Acela iaste, căruia Eu o *bucată* întinsă îi voiu da». Și întingînd pîinea, deade Iudei a lui Simon den Scariot. Și după ce luo bucata, atuncea întră întru el Satana” (NT, Ioan 13:26-27). În NT apare și sensul figurat de ‘hrană nepieritoare pentru suflet’: „Lucrați nu pentru bucate pieitoare, ce pentru *bucate* ce rămîn în viața veacilor” (Ioan 6:27). Un alt sens simbolic e acela de ‘hrană spirituală pentru cei formați întru credință, în opoziție cu *lapte*, care este destinat începătorilor’: „Cu lapte hrăniiu pre voi, nu cu *bucate*, ce încă nu puteați ce nece acum încă puteți” (NT 1Cor. 3:2); „Ce cînd va să dea omului lapte, elu e neștiutoriu în cuvîntul dereptăției, elu e un cocon tînăr. Iară adevăraților elu le trebuiaste *bucate* tari, ce pentru învățatura elu le-au dat minte ce întru despărțeală și al binelui și al răului” (LA, Evr. 5:13-14).

3.5. CARNE – subst. moștenit din lat. *caro, carnis* (DLR, *s.v.*). În corpusul nostru are următoarele atestări: *carne* (LA 5, NT3, B 1), *carnea* (LA 2, NT 8), *carneei* (NT 1). Sensul este de ‘hrană care nu respectă canoanele postului’: „Bine e a nu mîncă *carne* și a nu bea vin, nece a face ceva întru caria să poticneaste fratele tău, au să sminteaste sau slăbeaste” (NT, Rom. 14:21). Un alt sens este acela de ‘intrupare’: „Mă și vedeți, că duhul *carne* și oase n-are, precum Mă vedeți pre Mine” (B, Ioan 24:39).

3.6. DULCE – adj. moștenit din lat. *dulcis* (DLR, *s.v.*). În LA apare de 67 de ori; intră deseori în compuse de tipul: *dulce-vesti*, *dulce-mărturisî*, *dulce-dărui*, *dulce-vestirea*, *dulce-dăruire*, *dulce-dăruiind*, *dulce-răbdare* (NT 6, B 3, T 29); cuvinte înrudite: *dulceața* (LA 36, NT 1), *dulceață* (LA 36, B 6). În T este folosit de obicei cu sens figurat: „Ce mă grăiești *dulce*? Nimea nu e *dulce*, numai Dumnezeu unul” (T, Mat. 9:17). În acest verset, în loc de *dulce*, în NT apare termenul *bun* (Fig. 3); la fel în B.

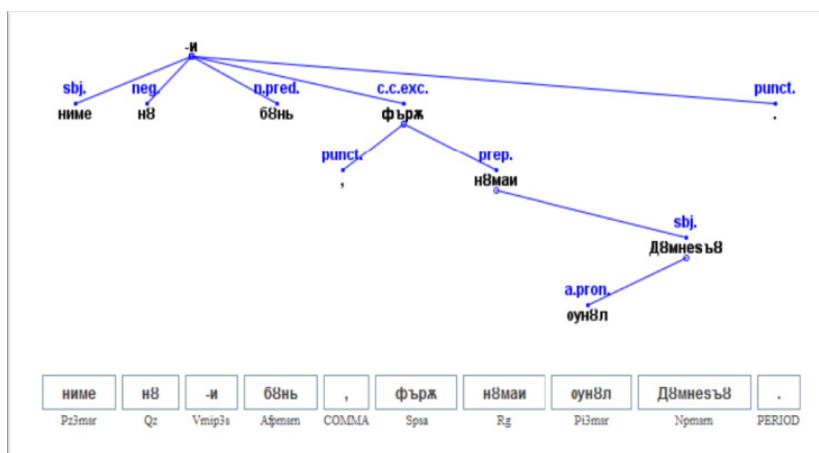


Fig. 3. Arbore adnotat morfo-sintactic, cu caractere chirilice.

3.7. FĂRÂMĂ – subst., element de substrat; cf. alb. *thërrime* (Ionescu 1985, p. 74). În textele studiate este atestat de mai multe ori: *fărîme* (T 7, B 2, NT 3), *fărîmele* (T 3, B 2, NT 3), *fărîmituri* (B 2, NT 1). Sensul este de ‘aliment care prisosește după ce necesitatea hrănirii este satisfăcută; belșug’. Sensul ‘hrană spirituală’ apare în relatarea despre înmulțirea pâinilor, când se precizează numărul de coșuri cu fărâme care au rămas în urma ospățului.

3.8. HRANĂ – subst., din vsl. *chrana* (DLR, *s.v.*). În textele studiate are următoarele ocurențe: *brană* (LA 7, NT 5, B 6, T 3), *brana* (NT 5, B 5), *brane* (B 1), *brancei* (T 1). Acolo unde în T apare termenul *brană* cu sensul propriu de ‘satisfacere a necesităților trupești’ (‘Pustiu easte locul și ceasul amu treace; lasă gloatele să margă împrejurul satelor să cumpere lor *brană*’, Mat. 14:15), în NT găsim termenul *bucate*. Termenul *brană* apare și cu sensul de ‘hrană spirituală’; din această perspectivă, este interesantă opoziția cu *laptele*, care este hrana pruncilor, adică a învățăceilor sau a ucenicilor în credință (vezi și *supra*).

3.9. LAPTE – subst. moștenit din lat. *lac, lactis* (DLR, *s.v.*). Atestări: *lapte* (LA 4, NT 4, B 3), *laptele* (LA 1, NT 1, B 2). Termenul *lapte* ‘hrană pentru prunci’ se flă în opoziție cu *brană întemeiată* ‘hrană pentru cei puternici’ (vezi *supra*): ‘Pentru că datori fiind a fi învățători pentru vreamă, iarăși vă trebuie a învăța pre voi carele-s stihile începutului cuvintelor lui Dumnezeu și v-ați făcut trebuindu-vă *lapte*, și nu hrană întemeiată’ (B, Evr. 5:12-13).

3.10. LEGUME – subst. moștenit din lat. *legumen, -is* (DLR, *s.v.*). Atestări: *legumi* (B 1), *legumile* (B 2). Termenul este folosit cu sens generic pentru o serie de condimente care diferă puțin în textele studiate; acestea sunt și un simbol pentru lucrurile mărunte scoase în mod greșit în evidență de farisei, care neglijează esențialul: ‘Ce vai voao, fariseilor, că zeciuiț izma și ruta, și toate *legumile*, și daș judecata și dragostea lui Dumnezeu; acestea s-au căzut a le face, și acelea a nu lăsa’ (B, Mat. 23:23). Elementele vămuite de farisei sunt, în contextul respectiv, următoarele⁴: *rută* ‘plantă erbacee aromatică’ (cf. lat. *ruta*; în B), *izmă* (element de substrat; în B, T, NT), *cimbru* (cf. gr. *thymbra*; în B), *chemin* (cf. gr. *keimion*; în NT), *pătrunjel* (cf. lat. *petroselinum*; în T), *molotru* (cf. slavon. *молотръ*; în NT), *mărar* (element de substrat, cf. alb. *mëraj*; în B, NT), *mintă* (cf. vsl. *мѣта*, lat. *mentha*; în NT).

3.11. MIERE – subst. moștenit din lat. *mel, mellis* (DLR, *s.v.*). Apare în mai multe texte: ‘Era Ioan îmbrăcat cu peri de cămilă și briu de curea prespre mijlocul lui și mânca lăcuste și *miare* sălbatecă’ (NT, Marc. 1:6), ‘mîncarea lui era mugur și *miiare* sălbatecă’ (T, Marc. 1:6), ‘mînca vlăstări și *miare* sălbatică’ (B, Marc. 1:6).

3.12. PÂINE – subst. moștenit din lat. *panis* (DLR, *s.v.*). Atestări: *pîne* (LA 8, NT 37, B 33), *pînea* (NT 24, B 17, LA 5), *pînei* (NT 2), *pîmile* (B 1), *pîni* (NT 36, B 23), *pîinii* (NT 1, B 2), *pîinile* (NT 6, B 6). În T (1561) apare termenul regional sinonim *pîtă* (cf. bg. *пита*), cu 14 atestări în T și 3 în LA, unde este atestat frecvent *pîne*. Putem observa

⁴ După DLR, *s.v.*

cum, în decurs de câțiva ani, Coresi renunță la regionalismul *pită*, într-un continuu efort de a selecta termenii cu circulație generală.

Este vorba despre unul dintre cele mai importante simboluri ale religiei creștine, simbol al transsubstanțierii trupului divin în hrană spirituală, prin care credincioșii devin un singur trup unii cu ceilalți și cu divinitatea. Sensul acestui termen este primordial pentru religia creștină și exemplifică perfect traiectoria de la sensul propriu la cel figurat-simbolic, deoarece prin transsubstanțiere și cuminecare se trece de la satisfacerea nevoilor trupești la hrana sufletească; cf. „Și mâncând ei, Iuș Iișus *pîne*, o blagoslovi, frîmșe-o și deade lor și zise: «Luăți, mîncăți, așta e trupul Mieu»” (NT, Marc. 14:22).

3.13. ROD – subst., cf. vsl. **рoдъ** (DLR, *s.v.*). Ateștări: *rod* (LA 11, NT 34, B 29), *rodul* (La 3, NT 26, B 14, T 12), *rodului* (NT 2, B 1, T 1), *rodure* (NT 2), *roduri* (B 2), *rodurile* (B 2). Pe lângă așest cuvânt, care are preponderent sens simbolic ori măcar generic, apar și atestări din familia sa lexicală: *rodeăște* (NT 2, B 1), *rodi* (LA 1, NT 1, B 1, T 1), *roditoare* (NT 1).

Fragmentul următor este adnotat semantic în Fig. 4: „Dară altele căzură în pămînt bun și deaderă *rod*: una, o sută, alta, șase zeci, iară alta treizeci” (NT, Mat. 13:8). Textul citat prezintă mai multe tipuri de elipse. În primul rând, nu apare cuvântul *sămânță*, care este un simbol atât pentru cuvântul propovăduit, cât și pentru persoana care se convertește și care va converti pe alții sau va avea urmași. În al doilea rând, nu este repetată locuțiunea verbală *a da rod*, pe care am adnotat-o transferând sensul ei (prin eticheta EQVH⁵) la elemente relaționale care devin regente pentru perechile următoare de subiect – obiect, sau, în adnotare semantică, existent – rezultat.

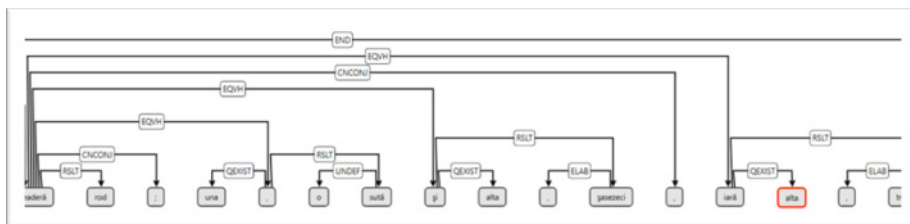


Fig. 4. Elipse. Mat. 13:8 (NT) în adnotare semantică.

Alt sens al cuvântului este cel de ‘faptă prin care se probează calitatea unui aspirant la viața de veci’ (pomul bun dă roade bune, cel care nu face roade nu este folositor, deci va fi tăiat).

⁵ Am introdus simbolurile EQVH (echivalent cu headul său, ținând locul lui), EQVZ și EQVP, ultimele două adnotând echivalența cu headul și schimbarea polarității pozitive în negativă și invers.

3.14. STRUGURI – subst., element de substrat (DLR, *s.v.*). Atestări: *struguri* (NT 2, B 2, T2), *strugurii* (NT 1, B 1); cf. „Au doară vor culeage den spini *struguri* au den scai smochine?” (NT Mat. 7:16, T Mat. 7:16). Acest cuvânt nu este doar o particularizare a genericului *rod*, ci, făcând referire la rodul viței din care se face vinul, element participant alături de pâine la transsubstanțiere, devine extrem de important. Via este și simbol al lumii, împreună cu viașii (poporul ales, cu care a fost semnat un legământ) care nu dau rodul ei stăpânului și îlucid pe fiul acestuia. Rodul *viței de vie* devine element unic pentru a dobândi „viața de veci”, cuvinte compuse, înrudite etimologic: „Iară zic voao că nu voiu bea de acum dentru acest rod de *vie* pînă la zua aceeaia carea voiu fi cu voi de nou întru *împărăția tatălui meu*” (NT, Mat. 26:29).

3.15. VIN – subst. moștenit din lat. *vinum* (DLR, *s.v.*). Atestări: *vin* (LA 5, NT 22, B 32, T 24), *vinul* (NT 18, B 16, T 3), *vinulu* (T 2), *vinului* (NT 1, B 3). Latura negativă a simbolului este prezentă alături de cea pozitivă: „Și nu vă îmbătați de *vin*, întru el iaste curvie, ce să fiți plin în Duh” (LA, Efes. 5:18); „Bea-va și acest din *vinul* mîniei lui Dumnezeu, carele s-au împlut în păharul mîniei Lui” (NT, Apoc. 14:10). Vinul *nou* este un simbol al *Noului Testament*: „Neci bagă *vinul nou* în foii vechi; că almintrea să sparg foii și *vinul* să varsă, și foii pier” (NT, Mat. 9:17). Transformarea apei în vin la nunta din Caana Galileei (nunta fiind și ea un nou legământ) este prima din seria minunilor christice. Relația dintre Iisus și biserica sa este văzută în multe parabole ca o nuntă: „Hristos întoarse apa în *vin* și făcu începiniia seamnelor Lui” (NT, Ioan, 2:11, rezumat inițial).

3.16. OLOI – subst., cf. magh. *olaj* (DLR, *s.v.*). Atestări: *oloi* (NT 2), *oloiul* (NT 1), *uloiu* (NT 5), *uloiul* (NT 2). Elementul este asociat deseori cu vinul: „Iară vinul și *oloiul* să nu-l strici” (NT Apoc. 6:6). Această combinație are puteri vindecătoare sau are capacitatea de a înlesni accesul în lumea de dincolo: „Și să apropie de el și legă ranele, turnînd în iale *uloiu* și vin; puindu-l pre el pre al lui dobitoc, duse-l în sălașul oaspeților și-i purtă căștiga” (NT, Luc. 10:34).

3.17. UNT – subst. moștenit din lat. *unctum* (DLR, *s.v.*). Atestări ale termenului individual, precum și în derivate sau în compuse: *unt* (LA 2, NT 1, B 1, T 6), *untu* (B 1, T 1), *untul* (B 1, T 1), *untdelemn* (B 7), *untdelemnul* (B 1), *untdelemnului* (B 1), *unsoare* (NT 10, B 4), *unsoarea* (NT 4), *unsoari* (B 3). Se observă că selecția sinonimelor variază în textele studiate. Ungerea rituală este un simbol al veseliei, bogăției și norocului: „Te-au uns Dumnezeu, Domnul Tău cu *unt* de veselie mai vârtos decît soții Tăi” (LA, Evr. 1:12). Prin ungere se exprimă respectul pentru musafir sau venerația pentru făcătorul de bine. De asemenea, untdelemnul este folosit în cadrul ritualurilor de înmormântare.

4. Concluzii

Referitor la terminologia hranei din textele studiate, mulți dintre termenii analizați au originea în substrat; alții sunt moșteniți din latină sau, mai rar, aparțin influenței slave. Aceștia apar atât cu sens propriu (hrana necesară trupului), cât și cu sens

figurat (hrana spirituală). După cum observăm în materialul analizat, trecerea de sensul propriu la cel figurat poate fi urmărită mai ales în parabolele biblice.

În ce privește studiul limbii române vechi, credem că acesta ar trebui să se poată efectua atât pe baza textelor transliterate cu alfabet latin, cât și pornind direct de la textele scrise cu alfabet chirilic (ca în cazul limbii grecești sau al limbii germane scrise cu caractere gotice). În acest sens, vom continua să îmbunătățim procesarea cât mai fidelă a cărților vechi.

Cât privește editarea unor texte cu apropierea textului de limba contemporană, pentru uzul nespecialiștilor în textele vechi sacre, nu ne putem opri la acest lucru atunci când vorbim despre moștenirea culturală a primelor secole de scriere în limba română. Instrumentele de procesare actuale includ în lexicoanele lor mai ales date despre limba română contemporană standard; prin obținerea clonelor în stare să proceseze limba veche în foma originală se va obține, implicit, capacitatea acestora de a procesa formele care conțin actualizări.

Corpusul RoDia este accesibil pe site-ul Facultății de Informatică și poate fi consultat. Cu programele Treeannotator (Fig. 3) sau Treebank Annotator (Fig. 2, 4) se pot vizualiza arborii semantici sau sintactici pe care îi conține corpusul. De asemenea, cu ajutorul programelor LUCON sau NOTEPAD++, se pot efectua căutări în corpusul în format XML folosit aici; acestea permit căutarea nu numai a cuvintelor, ci și a anumitor etichete de adnotare; de exemplu, *postag="Vmm-2p--l"* este o etichetă morfologică cu care am adnotat un fenomen specific limbii române vechi, i.e. imperativul format de la infinitivul lung (ex. *gândireți, gurarareți, protivireți, grăireți, închipuireți, fireți* etc.).

Bibliografie

A. Izvoare și lucrări de referință

B = *Biblia 1688*, în *Monumenta linguae Dacoromanorum. Biblia 1688, Pars XXIV: Novum Testamentum II: Actus Apostolorum, Pauli Epistolae, Catholicae Epistolae, Iudae, Apocalypsis Ioannis*, Eugen Munteanu (coord.), Anca Bibiri, Iosif Camară, Ana Catană-Spenchiu, Claudiu Coman, Emanuel Conțac, Marius Cruțeru, Ana-Maria Gînsac, Octavian Gordon, Gabriela Haja, Maria Moruz, Mihail Neamțu, Mihaela Paraschiv, Alina Pricop, Ovidiu Sferlea, Elena Tamba, Wilhelm Tauwinkl, Mădălina Ungureanu, Alin Vara, Mihai Vladimirescu. Editor: Ana Catană-Spenchiu, Iași, Editura Universității „Alexandru Ioan Cuza”, 2015.

CUV. B = *Cuvente den bătrâni*, tom. I: *Limba română vorbită între 1550-1600*, studiu paleografico-linguistic de B. Petriceicu Hasdeu; tom. II: *Cărțile populare ale românilor în secolul al XVI-lea* [...]; tom. III: *Istoria limbei române*, partea I: *Principie de lingvistică*. Ediție îngrijită, studiu introductiv și note de G. Mihăilă, București, Editura Didactică și Pedagogică, vol. I-III, 1983-1987.

DEL R = *Dicționarul etimologic al limbii române*, vol. I (A-B), București, Editura Academiei Române, 2010.

DLR = *Dicționarul limbii române*, ediție anastatică în 19 vol., după *Dicționarul Academiei* (DA) și *Dicționarul limbii române* (DLR), București, Editura Academiei Române, 2010.

- LA = Coresi, *Lucrul Apostolesc. Apostolul*, Brașov, 1563. În *Texte de limbă din secolul al XVI-lea*, vol. IV, reproduse în fascimile, îngrijite de Ioan Bianu, București, Tiparul „Cultura Națională”, 1930.
- NT = *Noul Testament*, tipărit pentru prima dată în limba română la 1648 de către Simion Ștefan, mitropolitul Transilvaniei, reeditat după 340 de ani, Alba Iulia, Editura Episcopiei Ortodoxe Române a Alba Iuliei, 1988.
- T = Coresi, *Tetraevanghelul tipărit de Coresi (Brașov, 1560-1561) comparat cu Evangheliarul lui Radu de la Mănăești, 1574*, ediție alcătuită de Florica Dimitrescu, București, Editura Academiei Române, 1963.

B. Literatură secundară

- Brâncuș, Grigore, *Vocabularul autohton al limbii române*, București, Editura Științifică și Enciclopedică, 1983.
- Cojocaru, Svetlana/ Colesnicov, Alexander/ Malahova, Ludmila, *Digitization of Old Romanian Texts Printed in the Cyrillic Script*, în *Proceedings of International Conference on Digital Access to Textual Cultural Heritage*, ACM Publisher, 2017, p. 143-148.
- Ionescu, Adriana, *Lexicul românesc de proveniență autohtonă în textele din secolele al XVI-lea și al XVIII-lea*, București, TUB, 1985.
- Iorga, Nicolae, *Istoria bisericii românești și a vieții religioase a românilor*, vol. I, Vălenii-de-Munte, Tipografia „Neamul Românesc”, 1908.
- Mărănduc, Cătălina/ Simionescu, Radu/ Cristea, Dan, *Hybrid POS-tagger for Old Romanian*, în *Proceedings of the International Conference on Computational Linguistics and Intelligent Text Processing CILing*, Budapest, 2017 (în curs de publicare).
- Mihăilă, Gheorghe, *Observații asupra manuscrisului slavo-român al popii Bratu*, în *Studii de limbă literară și filologie*, vol. II, 1972, p. 301-345.