

NADIA LUIZA CRUCERU

**NATURALEȚEA LIMBAJULUI
ARTIFICIAL**

**STUDIU DESPRE TRADUCEREA
AUTOMATĂ A GRUPULUI VERBAL**

NATURALEȚEA LIMBAJULUI ARTIFICIAL

STUDIU DESPRE TRADUCEREA AUTOMATĂ A GRUPULUI VERBAL

Copyright © 2021
Autor: NADIA LUIZA CRUCERU

Toate drepturile rezervate.

ISBN 978-606-9734-00-1

Editura Evomind, 2021

<https://evomind.org/>

Rezumat

Funcția specifică a limbii este de a servi societatea ca mijloc de comunicare, ca mijloc pentru schimbul de idei și informații între oameni. Dacă din punct de vedere tehnologic pot exista popoare primitive, nu aceeași realitate este valabilă la nivelul limbajului: antropologii au fost în repetate rânduri uimiți de complexitatea și gradul de abstractizare al limbilor societăților primitive ca tehnologie. De asemenea, limbajul vorbitorilor needucați este sistematic și condus de reguli, chiar dacă regulile aparțin altor dialecte decât cel standard.

Fiecare individ uman posedă un algoritm cu abilitatea de a genera un număr infinit de propoziții, fiecare dintre ele corespunzătoare unei idei distincte [77]. De exemplu, informația pe care o deținem despre gramatica limbii române sau a limbii engleze conține reguli de tipul “O propoziție poate fi compusă dintr-un grup nominal (subiectul) și un grup verbal (predicatul)”;

“Un grup verbal poate fi realizat dintr-un verb și un grup nominal (obiectul)”.

Gramatica se bazează pe reguli ce alcătuiesc subsisteme. Cel mai important subsistem este *sintaxa*, al cărei obiect de studiu îl reprezintă ansamblul regulilor privitoare la îmbinarea cuvintelor și relațiile dintre ele în propoziții și fraze.

Al doilea subsistem al gramaticii este morfologia, în care cuvintele sau părți din cuvinte (afixe) sunt asamblate pentru a produce cuvinte complexe.

Din punct de vedere sintactic, orice traducere se confruntă cu trei tipuri de probleme:

- i. definirea și recunoașterea tiparelor sintactice din limba sursă și a relațiilor care le compun într-un întreg acceptabil formal;
- ii. definirea și recunoașterea structurilor sintactice și a relațiilor echivalente din limba țintă;
- iii. corelarea celor două grupe de fenomene (procedurile de transfer dinspre limba sursă spre limba țintă).

În lucrarea **Naturaletăa limbajului artificial - studiu despre traducerea automată a GV**, analizez cele trei grupe de probleme, focalizându-mă asupra structurii sintactice a grupului verbal din limbile română (limba sursă) și engleză (limba țintă).

Etapele de lucru în realizarea acestei lucrări au cuprins citirea de documentație pentru o înțelegere mai bună a domeniului Prelucrării Automate a Limbajului Natural, definirea grupului verbal și alegerea unei modalități de reprezentare sintactică a constituenților grupului verbal, definirea și crearea tehnologiei pentru descrierea modulelor ce cuprind studiul de traducere automată a grupului.

Din punct de vedere teoretic, lucrarea își propune să prezinte, pentru limba română, o abordare personală a morfologiei paradigmatică, ținând cont de două tipuri de informații: restricția de nominativ-acuzativ plural, nearticulat, considerată identificator de paradigmă, și operațiile de *alternanță*, *adunare sau scădere de șiruri*. Metoda propusă a fost pusă în aplicare în cadrul modulului de deflexionare nominală și verbală, unul dintre modulele ce compun studiul de traducere automată.

La nivelul dezambiguizării am abordat interactiv modulul de preprocesare a textelor TTL, creat și dezvoltat la Institutul de Cercetare pentru Inteligența Artificială, astfel încât pentru lexemele considerate ambigue prin numărul de etichete morfosintactice este apelat programul TTL care reține pentru fiecare ocurența lema și eticheta morfosintactică specifice.

Flexiunea nominală și verbală pentru limba engleză a fost realizată prin intermediul regulilor flexionare și pe baza opoziției dintre forme regulate vs. forme neregulate. Formele neregulate sunt înregistrate în baza de date a aplicației, nefiind necesară o procedură de flexionare, ci doar una declarativă, strict de memorare. În schimb, formele regulate sunt clasificate pe baza comportamentului morfosintactic, iar regulile au în vedere șabloane de flexiune tratate procedural.

Elaborarea secțiunii de sintaxă a grupului verbal pentru limbile română și engleză are la bază o etapă de cercetare în care am definit constituenții grupului verbal, caracterul obligatoriu sau facultativ de ocurență în cadrul grupul verbal și diferite tipare verbale. Definirea determinanților verbali a fost urmată de alegerea unui instrument pentru reprezentarea grafică a arborilor sintactici, selecția oprindu-se asupra programului

Linguistic Tree Constructor, preluat din catalogul de programe al site-ului www.sil.org. Implementarea ulterioară în mediul de lucru Delphi a regulilor de identificare a grupului verbal, create în etapa de cercetare și realizarea aplicației AGV1.0 (Analizor de Grup Verbal 1.0) au constituit etapa finală a studiului de traducere automată a grupului verbal.

Cuprins

	Listă de figuri	3
	Listă de tabele	5
1	Introducere	6
	1.1 Limbajul- interacțiune între lexicon și gramatică	6
	1.2 Definirea și structura grupului verbal	10
	1.2.1 Definirea grupului verbal.....	10
	1.2.2 Structura grupului verbal	13
2	Deflexionarea nominală și verbală în limba română	16
	2.1 Introducere	16
	2.2 Paradigmele nominale în limba română	23
	2.2.1 Prima metodă de deflexionare nominală.....	24
	2.2.2 Algoritmul de deflexionare nominală	28
	2.2.3 A doua metodă de deflexionare nominală	31
	2.3 Paradigmele verbale în limba română.....	33
	2.3.1 Prezentarea cadrului de lucru.....	33
	2.3.2 Obținerea formei verbale bază.....	45
3	Flexiunea nominală și verbală în limba engleză	50
	3.1 Flexiunea nominală în limba engleză	50
	3.1.1 Problematizare	50
	3.1.2 Schema logică de reprezentare a flexiunii nominale	53
	3.1.3 Reguli de flexionare a numelui comun.	54
	3.2 Conjugarea verbului în limba engleză	55
	3.2.1 Introducere	55
	3.2.2 Schema logică de conjugare a verbului	58
	3.2.3 Obținerea formei verbale conjugate.....	59
	3.2.4 Grupele de verbe din limba engleză, diateza activă.....	61
	3.3 Descrierea implementării flexiunii nominale și verbale	66

3.3.1	Conjugarea verbului.....	66
3.3.2	Flexiunea nominală.....	69
4	Modulul de preprocesare a textelor TTL.....	72
	4.1 Prezentare generală.....	72
	4.2 Utilizarea modulului de preprocesare a textelor TTL.....	75
5	Sintaxa grupului verbal pentru limbile română și engleză	79
	5.1 Tipuri de gramatici generative	79
	5.2 Linguistic Tree Constructor	83
	5.3 Etapa de cercetare. Verbele bivalente	86
	5.3.1 Detalierea structurilor verbale.....	88
	5.4 Etapa de cercetare. Verbele trivalente.....	100
	5.4.1 Obiectul direct și obiectul indirect	100
	5.4.1.1 Dezvoltarea structurilor verbale cu obiect direct și obiect indirect	102
	5.4.2 Structuri verbale cu obiect direct și complement secundar.....	107
	5.4.2.1 Dezvoltarea structurilor verbale cu obiect direct și complement secundar.....	108
	5.4.3 Structuri verbale cu obiect direct și complementul predicativ al obiectului.....	111
	5.4.3.1 Dezvoltarea structurilor verbale cu obiect direct și complementul predicativ al obiectului.....	111
	5.4.4 Structuri verbale cu obiect direct și obiect prepozițional.....	115
	5.4.4.1 Dezvoltarea structurilor verbale cu obiect direct și obiect prepozițional	116
	5.5 Descrierea implementării analizei grupului verbal	130
6	Concluzii	133
	Bibliografie	135

Listă de figuri

- 1.1 Proprietățile de selecție ale verbului
- 2.1 Exemple de reguli tipice pentru morfologia pe două niveluri
- 2.2 Paradigmă verbală globală
- 2.3 Paradigmă verbală parțială
- 2.4 Paradigmă nominală pentru substantive masculine cu terminațiile {și, ști}
- 2.5 Paradigmă flexionară pentru substantivele feminine cu terminațiile *ete, ețe, epe, erze, etre, este*
- 2.6 Pașii de analiză a formei *fetele*
- 2.7 Identificarea șirului *le* ca restricție pentru forma *zilele*
- 2.8 Activarea restricției *le* pentru forma *fetele*; fără validarea formei finale
- 2.9 Etapele procesului de deflexionare nominală. Prima metodă
- 2.10 Definirea obiectului *Restrictia*
- 2.11 Situație conflictuală cu două restricții potențiale pentru un cuvânt
- 2.12 Etapele procesului de deflexionare nominală. A doua metodă
- 2.13 Schema logică de analiză a verbului
- 3.1 Ilustrare simplificată a teoriei Cuvinte-și-Reguli
- 3.2 Schema logică de reprezentare a flexiunii nominale
- 3.3 Paradigma verbelor lexicale
- 3.4 Schema logică de reprezentare a conjugării verbului
- 4.1 Compuși românești ca unități lexicale
- 4.2 Prefixe (LEFTSPLIT) și sufixe (RIGHTSPLIT) care trebuie separate în română
- 5.1 Fereastra de deschidere a programului LTC
- 5.2. Deschiderea arborilor sintactici pe verticală (prima figură) și pe orizontală (a doua figură) pentru exemplul 1
- 5.3 Schema generală de reprezentare a grupului verbal cu nucleu bivalent
- 5.4 Deschiderea arborilor sintactici pe orizontală pentru exemplul 2
- 5.5 Deschiderea arborilor sintactici pe orizontală pentru exemplul 3
- 5.6 Arborii sintactici asociați exemplului 4

- 5.7 Arborii sintactici asociați exemplului 5
- 5.8 Arborii sintactici asociați exemplului 6
- 5.9 Arborii sintactici asociați exemplului 7
- 5.10 Arborii sintactici asociați exemplului 8
- 5.11 Deschiderea arborilor sintactici pe orizontală pentru exemplul 9
- 5.12 Structura verbală cu aritate 3 de tipul (verb tranzitiv, obiect direct și obiect indirect)
- 5.13 Deschiderea arborilor sintactici pe orizontală pentru exemplul 10
- 5.14 Deschiderea arborilor sintactici pe orizontală pentru exemplul 11
- 5.15 Deschiderea arborilor sintactici pe orizontală pentru exemplul 12
- 5.16 Deschiderea arborilor sintactici pe orizontală pentru exemplul 13
- 5.17 Structuri verbale cu obiect direct și complement secundar
- 5.18 Arborii sintactici asociați propoziției cu verb la modul imperativ
- 5.19 Arborii sintactici asociați structurii verbale cu obiect direct și complement secundar
- 5.20 Structuri verbale cu obiect direct și complement predicativ al obiectului
- 5.21 Deschiderea arborilor sintactici pe orizontală pentru exemplul 14
- 5.22 Deschiderea arborilor sintactici pe orizontală pentru exemplul 15
- 5.23 Deschiderea arborilor sintactici pe orizontală pentru exemplul 16
- 5.24 Deschiderea arborilor sintactici pe orizontală pentru exemplul 17
- 5.25 Deschiderea arborilor sintactici pe orizontală pentru exemplul 18
- 5.26 Deschiderea arborilor sintactici pe orizontală pentru exemplul 19
- 5.27 Deschiderea arborilor sintactici pe orizontală pentru exemplul 20
- 5.28 Deschiderea arborilor sintactici pe orizontală pentru exemplul 21
- 5.29 Arborii sintactici asociați exemplului 22
- 5.30 Arborii sintactici asociați exemplului 23
- 5.31 Arborii sintactici asociați exemplului 24
- 5.32 Deschiderea arborilor sintactici pe orizontală pentru exemplul 25

Listă de tabele

- 2.1 Comportamentul flexionar al substantivelor *urs* și *artilerist*
- 2.2 Comportamentul flexionar al substantivelor ce respectă restricția *ete*, *ețe*, *epe*, *erze*, *etre*, *este*
- 2.3 Clasă generală de terminații pentru verbele de conjugarea I (ConjI.CGT)
- 2.4 Clase parțiale de terminații
- 2.5 Alternanțele fonetice aplicate celei de a doua grupe de verbe de conjugarea I
- 2.6 A treia grupă de verbe de conjugarea I (ConjI.V3)
- 2.7 Terminații pentru prima grupă de verbe de conjugarea a II-a
- 2.8 Terminațiile pentru grupele de verbe 2-4, conjugarea a II-a (ConjII.V2, ConjII.V3, ConjII.V4)
- 2.9 Clasă parțială de terminații
- 2.10 Clasele parțiale de terminații ConjIII.CPT 2, ConjIII.CPT 3, ConjIII.CPT 4
- 3.1. Conjugarea verbelor din grupa de terminații T1 la modul indicativ
- 3.2 Modul indicativ, forma afirmativă: verb_{afirmativ}
- 3.3 Modul indicativ, forma negativă: verb_{negativ}

Capitolul 1

Introducere

1.1 Limbajul- interacțiune între lexicon și gramatică

În lingvistica secolului al XX-lea, limbajul a fost interpretat ca „rețea de relații” (Hjelmslev) sau ca „structură sintactică” corespunzătoare „competenței lingvistice a unui vorbitor ascultător ideal” (Chomsky). Hjelmslev așază la baza conceptului său de limbaj faptul că: „Raporturile (...) sunt primare, în sensul că existența obiectului e considerată ca un rezultat al lor”. Altfel spus: „o totalitate nu constă din obiecte, ci din raporturi și (...) ceea ce are o existență în sensul științific al cuvântului nu e substanța acelei totalități, ci numai raporturile ei interne și externe”¹. El elimină din definiția semnului, ca relație între forma expresiei și forma conținutului, „substanța” semantică a limbajului, tocmai pentru a se ajunge la „structura algebrică” a limbii considerată ca fiind obiectul științei limbajului [55].

Această abordare ignoră însă tocmai faptul esențial, anume că: „în limbaj, totul este semantic: gramatica nu e mai puțin [semantică] decât lexiconul, cuvântul în general și limbile nu sunt astfel mai puțin [semantice] decât discursul” și de aceea „a vorbi despre semantică echivalează, în consecință, cu a vorbi despre întreaga lingvistică”².

Cel mai remarcabil aspect al limbajului este puterea sa expresivă: capacitatea de a exprima un număr nelimitat de idei de la o persoană la alta prin intermediul unui flux structurat de sunete. Cercetările întreprinse de lingvistul Otto Ducháček se sprijină tocmai pe această primă modalitate de interpretare a raporturilor cuvânt – concept – lucru. Concepția sa despre limbă „en tant que moyen de communication” are, în această perspectivă, ca punct de plecare faptul că, în ultimă instanță, fenomenul lingvistic se reduce la stabilirea unei „legături” între „anumite percepții acustice (eventual grafice) cu anumite noțiuni care reflectă realități exterioare”. În plus, „scopul principal al limbii este de a permite indivizilor de a se înțelege unii cu alții”³.

¹ L. Hjelmslev, **Preliminarii la o teorie a limbii**, trad. rom. D. Copceag, București, p. 41

² E. Coseriu, **Pour et contre l'analyse sémiologique**, Tokio, p. 137

³ Ducháček, 1967 – **Précis de sémantique française**, Brno, p. 20

Un limbaj nu este doar un cod conceput pentru a potrivi eficient sunetul cu înțelesul. El se bazează pe două principii fundamentale, cuvintele și gramatica.

Primul principiu are în vedere lexiconul mental, o listă memorizată și finită de cuvinte. După Ferdinand de Saussure, cuvântul este un semn alcătuit din două laturi: conceptul (noțiunea), care e latura semnificată- *le signifié*, și imaginea acustică a complexului sonor, care are rolul de a semnifica acel concept; Saussure o numește *le signifiant*. El precizează: “Numim *semn* îmbinarea conceptului cu imaginea acustică; dar în vorbirea curentă acest termen desemnează de obicei numai imaginea acustică, de pildă un cuvânt (*arbore* etc.)”⁴

Relația dintre conținut și expresie formează conceptul lingvistic. De exemplu, în cazul conceptului lingvistic *arbore*, expresia sau forma sonoră a cuvântului este a-r-b-o-r-e, iar conținutul cuvântului este „faptul de a fi arbore”. Până la Wilhelm von Humboldt, conținuturile cuvintelor erau înțelese ca fiind elaborate anterior expresiei în mintea umană. Aceste conținuturi elaborate și stocate în mintea umană se considera că, în vorbire, ar fi exprimate în forma materială (expresia sau forma sonoră) specifică diverselor limbi. Înainte de Humboldt, teoretizările asupra limbajului s-au bazat pe ideea că există în mintea umană un fel de conținut universal (aceiași pentru toți oamenii și identic structurat în orice limbă), înainte de a fi exprimat în cuvinte.⁵

Limbajul maximizează avantajele diferite ale cuvintelor și regulilor prin considerarea acestora ca fiind manipulate fiecare de câte un sistem psihologic propriu. Există astfel un lexicon pentru cuvinte generale sau pentru idiomuri; mecanismul psihologic creat pentru a-l manipula este un anumit tip de memorie. Există, de asemenea, un sistem separat de reguli de gramatică pentru noile combinații de entități; mecanismul psihologic creat pentru el este procesarea simbolică. Această teorie care acoperă cuvintele și regulile poate fi verificată în ceea ce privește flexiunea regulată și neregulată. Cuvintele cu flexiune neregulată sunt memorizate într-un lexicon mental. În schimb, cuvintele regulate au un comportament flexionar predictabil și, prin urmare, supus unor reguli de transformare [74]. Un exemplu de flexiune regulată e dat de formele englezești de *Past Tense* precum “walk - walked”, “kiss - kissed”, etc. Aproape toate verbele din

⁴ F. de Saussure, **Curs de lingvistică generală**, Editura Polirom, Iași, 1998, p. 99

⁵ W. von Humboldt susține că nu există conținut elaborat mai întâi în minte care apoi să fie doar exprimat în cuvinte, idee contrazisă însă de mulți psiholingviști și teoreticieni actuali ai limbajului.

limba engleză sunt regulate, iar clasa lor este complet predictabilă. Productivitatea deschisă a șablonului regulat și predictabilitatea lui sugerează că formele temporale de *Past Tense* sunt generate printr-o regulă mentală, similară ca formă altor reguli de gramatică de tipul “pentru a forma *Past Tense*, adaugă sufixul *-ed* la radicalul verbal”.

Pentru evitarea unei posibile erori terminologice, conceptul *cuvânt* primește următoarele unități fundamentale: *morfemul* (unitate minimală dotată cu sens, de obicei obiect al morfologiei) și *sintagma* (orice combinație de morfeme), obiect al sintagmaticii sau al sintaxei⁶.

O propoziție are o structură ierarhică permițând exprimarea propozițiilor complexe constând din idei incluse în alte idei. Pentru a putea exprima legăturile complexe dintre obiectele lumii reale, cuvintele se grupează în cursul vorbirii, stabilind între ele anumite relații ce determină precis funcția fiecăruia într-un enunț dat. În literatura lingvistică, tipurile de relații sintactice sunt introduse astfel [12] :

- clasificarea lui L. Bloomfield, generată de criteriul distribuțional, prezintă relații de tip: coordonativ, subordonativ și exocentric, argumentând prin distribuția relațională: similară vs. diferită;

- tipurile de dependență existente între termenii oricărei relații sintactice, impuse de L. Hjelmslev [46]: determinare (dependență unilaterală obligatorie), interdependență (dependență bilaterală obligatorie), respectiv, constelație (dependență facultativă);

- distincția între „sintagmă” și „îmbinare de cuvinte” făcută de unii lingviști sovietici: sintagma este, în opinia lor, o noțiune mai mult semantică (și mai puțin sintactică);

- opțiunea lingvisticii americane descriptive pentru interpretarea relațiilor sintactice din interiorul propoziției ca îmbinări de constituenți;

- Gramatica Academiei propune definiția: „Cea mai mică unitate sintactică în interiorul căreia se poate stabili un raport sintactic este îmbinarea de cuvinte, constituită din cel puțin două cuvinte cu sens lexical deplin. Nu poate apărea de sine stătătoare, este lipsită de predicție și nu poate constitui o unitate de comunicare decât prin încadrarea ei într-o propoziție”. Mioara Avram⁷ reia parțial definiția dată de Gramatica Academiei, și

⁶ Valeria Guțu Romalo, Gramatica limbii române, EDP, București, 1973, p. 5.

⁷ Mioara Avram, Gramatica pentru toți, ediția a III-a, Editura Humanitas, București, 2001, p. 403

anume: îmbinarea (grupul) de cuvinte, deci sintagma, este un constituent al propoziției alcătuit din cel puțin două cuvinte cu sens lexical deplin și reprezentând cea mai mică unitate sintactică în interiorul căreia se poate stabili un raport sintactic.

- Din opiniile lingviștilor români, pentru definirea sintagmei reținem aspectele:

- grup de două părți de propoziție în raport de subordonare⁸;
- două sau mai multe cuvinte, cu sens unitar, aflate într-un raport de subordonare⁹.

Concluzionând, reținem faptul că în această lucrare voi considera sintagma ca unitatea semantico-sintactică formată din doi termeni (morfeme, cuvinte, construcții gramaticale, propoziții) și relația dintre ei.

Sintaxa implică, de asemenea, structura predicat – argument, o componentă a limbajului care codifică relațiile existente la nivelul unui grup de participanți. Pentru a înțelege o propoziție, nu este suficient a acorda atenție numai ordinii cuvintelor sau modului în care acestea sunt grupate; de importanță maximă în stabilirea semnificației este interesul pentru informația asociată predicatului care specifică modul în care argumentele sunt așezate în propoziție. De exemplu, în propozițiile “Omul se teme de câine” și “Omul sperie câinele”, cuvântul “omul” este subiect în ambele cazuri, dar rolul lui semantic diferă în funcție de modul în care verbul consideră că “subiectul este cel care experimentează teama” sau “subiectul este cel care cauzează teama”. În prima propoziție omul experimentează teama; în a doua el o produce.

Sintaxa are mai multe posibilități de exprimare a aceleiași idei decât morfologia. De exemplu, pentru a exprima categoria persoanei I plural, limba română folosește desinența *-m* indiferent de conjugarea verbului, de timp sau mod: “luptăm”, “vedeam”, “dusesem”, “vorbirăm”. Dar pentru a reda în sintaxă ideea de cauză, vorbitorul are posibilitatea să aleagă între mai multe forme de exprimare: “deoarece (fiindcă, pentru că, de vreme ce) e frig, m-am îmbrăcat gros”; “e frig, m-am îmbrăcat gros”; “fiind frig, m-am îmbrăcat gros”.

Spre deosebire de celelalte domenii lingvistice, sintaxa impune o cunoaștere amplă și o înțelegere mai profundă a limbii. Nevoia firească a creării și a înțelegerii mesajului unui text a determinat studierea mecanismelor de funcționare a limbii, fiind

⁸ Sorin Stati, Elemente de analiză sintactică, Editura Didactică și Pedagogică, București, 1972.

⁹ Gheorghe Bulgăr, Limba română. Fonetică, lexic, morfologie, sintaxă, stilistică, Editura Vox, București, 1995

îndreptățită observația că sintaxa reprezintă nivelul superior al funcționării complete a limbii.

1.2 Definirea și structura grupului verbal

1.2.1. Definirea grupului verbal

În Gramatica Academiei¹⁰ se consideră grup verbal o parte componentă a enunțului, organizată în jurul unui centru verbal (la moduri personale sau nepersonale) și alcătuită din verbul-centru și toți constituenții legați sintactic de acesta, respectiv termenii care satisfac o valență a nucleului verbal.¹¹ Definiția orientează atenția asupra constituirii unui grup verbal în jurul unui centru verbal nepredicativ, afirmație detaliată anterior de către Gabriela Pană Dindelegan în lucrarea **Sintaxa transformațională a grupului verbal în limba română** unde face o distincție clară între grupul verbal și cel predicativ: Grupul verbal, o noțiune categorială, „neutru din punctul de vedere al funcției, este organizat în jurul oricărui centru verbal, indiferent de predicativitatea sau nonpredicativitatea acestuia.” Grupul predicativ, noțiune funcțională, „legat obligatoriu de funcția predicativă, este organizat în jurul unui verb la mod personal.”

În *Dicționarul de lingvistică și fonetică*, realizat de către David Crystal, în 1980, pentru termenul *grup verbal* sunt citate două accepții¹². În mod tradițional, conceptul se referă la un grup de verbe care au împreună o singură funcție sintactică (de exemplu, „is coming”, „may be coming”, „get up to”). Pentru astfel de contexte, unul dintre verbe este verbul de bază sau lexical și domină celelalte elemente. Un verb urmat de o particulă non-verbală (similară prepoziției sau adverbului) este, în genere, referit ca un verb frazal. Conform celei de a doua accepții, în gramatica generativă, grupul verbal este echivalent cu întregul predicat al propoziției, ilustrând expansiunea acesteia ca o combinație minimală de grup nominal și grup verbal.

¹⁰ Academia Română, Institutul de Lingvistică “Iorgu Iordan- Al. Rosetti” (2005) **Gramatica Limbii Române**, vol. I și II, Editura Academiei Române, București.

¹¹ Op. cit., p. 47, vol. al II-lea

¹² Ibidem, pp. 374, 1980

Din perspectiva substructurilor grupului verbal, în [18] se specifică: subgrupul auxiliarului de timp (Tense Auxiliary SubGroup- TASG), subgrupul auxiliarului de pasiv (Passive Tense Auxiliary SubGroup- PTASG) și subgrupul modal (Modal Verbal SubGroup- ModVSG). TASG este cel mai frecvent și mai natural subgrup verbal, este nesaturat și nepredicațional, deoarece necesită un verb predicațional care să-i completeze sensul (ex.: „voi fi”, „aș fi”, „am fi”, „sunt”). PTASG are sensul de atribuire copulativă și necesită un verb nefinit, un substantiv sau un adjectiv pentru a constitui grupul verbal (ex.: „a fost” în „Floarea [i-a fost] dată mamei tale”.) ModVSG cere ca argument o altă predicție, de ordin unu sau doi și are ca nucleu semantic un verb modal (ex.: „pot”, „aș fi putut”, „ar fi trebuit”).

Rezultate diferite [81] au arătat necesitatea unei reprezentări mai organizate a relației dintre predicat și argumente, pe baza unor principii abstracte. Există, fără îndoială, o legătură între proprietățile semantice ale unui verb și cele sintactice, aceasta explicând de ce, dezambiguizând înțelesul unui lexem, vorbitorii completează modelul comunicațional, intuindu-i forma sintactică [21].

Fiecare predicat are (sau, cel puțin, se postulează că ar avea) o structură de argumente ce specifică numărul minim de participanți ceruți de verb pentru ca acțiunea sau starea descrisă de acesta să fie suficient de clară.

O structură de argumente a unui verb definește numărul și relațiile participanților necesari pentru ca o acțiune să fie complet definită. Proprietățile de selecție ale verbului determină tipul contextului sintactic în care acest element poate apărea. Verbul exemplificat în (1a) este intransitiv (nu admite complement direct); cel din propoziția (1b) este tranzitiv (poate primi în context unul sau două complemente directe); în propoziția (1c) este un verb care cere un complement prepozițional; în (1d) verbele admit un complement direct și unul prepozițional:

- a. Elefantul *aleargă* fericit.
- b. Copiii *iubesc* viața.
- c. Alex *se gândește* la cartea lui.
- d. El *amenință* politicianii că va alerta presa.
- e. Ea *compară* acest tablou cu cele pictate de Grigorescu.

Figura 1.1: Proprietățile de selecție ale verbului

Extinderea grupului verbal diferă în funcție de:

- a. trăsăturile de subcategorizare ale verbului-nucleu;
- b. forma de diateză și, respectiv, forma modală din paradigma verbului (existând diferențe între paradigmele verbului personal și ale celui nepersonal);
- c. satisfacerea tuturor valențelor verbului-nucleu sau numai a unora dintre ele.

În lucrările de sintaxă modernă, pentru desemnarea determinanților¹³ la proces se mai folosește și termenul de *actant*; ansamblul actanților care gravitează în jurul verbului formează structura actanțială a propoziției. Numărul actanților regizați de verb constituie unul dintre criteriile esențiale de clasificare a verbelor. Teoretic poate fi construit un anumit număr de modele valențiale caracteristice limbii române [44]:

- verbe avalente (0+verb+0): a ninge (*Ninge.*);
- verbe monovalente (admit un nominativ): a fugi (*Cățelul fuge.*);
- verbe bivalente (admit cazurile următoare: fie două nominative- a deveni (*El va deveni un arhitect.*) sau un nominativ și un acuzativ: a scrie (*Tu scrii niște versuri.*) sau un nominativ și un dativ: a mulțumi (*Noi i-am mulțumit.*);
- verbe trivalente (au trei valențe cazuale): a oferi (*Tu vei oferi prăjituri soției.*)
- verbe tetravalente (se pot combina cu patru cazuri distincte, dintre care unul este cazul X sau cazul neutru, direct sau adverbial, ocupat de substantive cu valoare adverbială și funcție de circumstanțial sau de element predicativ suplimentar: a dărui (*El ți-a dăruit florile mănunchi.*)

¹³ participanți activi sau pasivi la proces

1.2.2 Structura grupului verbal

Sistematizând, grupul verbal primește următoarea componență:

- A. În funcție de includerea adjuncturilor în structura grupului verbal:
1. Grup verbal constituit numai din verbul centru și actanții sau complementele sale: [Verb + Complemente];
 2. Grup verbal care include și Grupul Circumstanțial (Circ), reprezentat de subordonații facultativi: [[Verb + Complemente] + Circ].
- B. În funcție de includerea subiectului în structura grupului verbal:
1. Grup verbal cu subiect necuprins în componența sa; subiectul se așază la același nivel ierarhic cu verbul, fiind interpretat ca a doua parte principală de propoziție;
 2. Grup verbal în care subiectul este considerat complement de tip special, căruia verbul îi impune restricții de caz și de rol tematic;
 3. Grup verbal în care subiectul este cerut nu numai de verb, ci de întregul grup [Verb + Complemente].

Ținând cont de aceste considerente, extinderea grupului verbal se realizează potrivit particularităților semantico-sintactice ale verbului nucleu, formei personale și de diateză din paradigma verbului-centru și după actualizarea tuturor disponibilităților combinatorii ale verbului-nucleu sau numai a unora dintre ele. /±Obligativitatea/determinantului este o trăsătură sintactică inerentă, intrinsecă a fiecărui verb în afara oricărei condiționări contextuale și a oricărei actualizări [66]:

3.1 au statut obligatoriu determinanții care participă prin transfer semantic la configurația semantică a verbului;

3.2 au statut facultativ determinanții care nu influențează nucleul semantic definitoriu al verbului și nu participă la regulile de transfer semantic¹⁴.

Structurile de tip [Verb + Complemente] constituie, în organizarea grupului verbal, tiparul său sintactic central, celelalte componente subordonându-se acestei structuri, coezive semantic și sintactic.

Coeziunea semantică este “efectul constrângerilor de rol tematic impuse de verb complementelor și subiectului, verbul prin trăsăturile lui semantic matriciale,

¹⁴ cf. G. Pană Dindelegan, 1978, p. 593

determinând tiparul de structurare semantică a propoziției” (vol. 2, p. 52). Rezultă diverse scheme corespunzătoare unor clase verbale, invariante la modificările pe care nucleul verbal le-ar putea primi la o formă personală ori nonfinită:

- a. [Verb + Agent];
- b. [Verb + Pacient];
- c. [Verb + Agent + Pacient];
- d. [Verb + Agent + Pacient + Țintă];
- e. [Verb + Beneficiar + Temă];
- f. [Verb + Experimentator + Temă];
- g. [Verb + Experimentator + Locativ];
- h. [Verb + Pacient + Locativ];
- i. [Verb + Pacient + Țintă].

Coeziunea sintactică este asigurată de restricțiile cazuale și de prepoziții impuse de verb complementelor sale, aceste mărci de actanță având rolul de a lega complementele de verb, pe de o parte, respectiv, de a diferenția tipurile de complemente, pe de altă parte.

Subclasificarea verbului în funcție de obiectul direct ține de distincția *tranzitiv / intransitiv*. O definiție complexă a tranzitivității o găsim în Gramatica Academiei (vol. 1, p. 341) unde se consideră că sunt tranzitive acele verbe care cer sintactic și semantic poziția obiectului direct acoperită sau neacoperită în context, poziție în care figurează un nominal substantival cu formă de Nominativ-Acuzativ, dacă acest nominal nu este subiect (deci nu se acordă cu verbul) și nici nume predicativ (nu este dependent de numele-subiect).

Alte definiții ale tranzitivității existente în lucrările de gramatică românească aparțin lingviștilor Mioara Avram, Ștefania Popescu, Gh. Constantinescu-Dobridor, Iorgu Iordan, Gabriela Pană Dindelegan:

a. „După o trăsătură sintactică legată de sensul lexical, verbele se clasifică în *tranzitive și intransitive*. Sunt tranzitive verbele care pot primi un complement direct: *Bea apă. Coase un nasture* (când complementul direct nu este exprimat, verbele sunt tot tranzitive, dar folosite absolut: *Bea prea mult, A cusut toată ziua*). Sunt intransitive verbele care nu pot avea un complement direct: *Stau în București, Îți aparțin*.” [4]

b. „Verbele care pot avea un complement direct sînt tranzitive, iar cele care nu pot avea un complement direct sunt intransitive.” [79]

c. „Sunt tranzitive acele verbe care au capacitatea de a realiza în mod obișnuit relația cu obiect direct sau cu o completivă directă. Sunt intransitive acele verbe care nu posedă capacitatea de a governa un obiect direct sau o completivă directă.” [15].

d. „Relația verb – obiect, deci relația *verb – substantiv în acuzativ*, constituie criteriul altei clasificări a verbelor: clasificarea în verbe *tranzitive și intransitive*. Sunt tranzitive verbele care au capacitatea de a realiza această relație, deci a governa un substantiv (pronume) în acuzativ, și intransitive cele caracterizate negativ sub aspectul acestei relații.” [47].

e. Gabriela Pană Dindelegan delimitează verbele tranzitive pe baza următoarelor teste formale ale tranzitivității: a) construcția cu un complement direct realizabil printr-un clitic pronominal cu forma de acuzativ; b) dublarea obiectului; c) pasivizarea.

Lista determinărilor obligatorii dominate de verb cuprinde, în afară de subiect¹⁵, complementul direct, complementul secundar, complementul indirect, complementul prepozițional, numele predicativ, complementul predicativ al obiectului, circumstanțialul (cerut obligatoriu de unele verbe).

¹⁵ Subiectul este acceptat de **Gramatica Academiei** drept un complement de tip special, dependent sub multe aspecte de verb, dar impunând el însuși restricții de acord, și având calitatea de complement extern

Capitolul 2

Deflexionarea nominală și verbală în limba română

Limbajele naturale realizează, prin secvențe, niveluri lingvistice structurate: morfologic, sintactic, semantic, pragmatic. Totuși, această idee este o reprezentare foarte simplificată. Întreg procesul care generează sistemul lingvistic construiește o punte între niveluri, fiecare contribuind la dezvoltarea și înțelegerea celuilalt (entitățile sintactice au rădăcină în nivelul morfologic, iar cele pragmatice în nivelul sintactico-semantic).

Secțiunea de morfologie a acestei lucrări propune o abordare analitică a limbii române utilizând paradigmele nominale și verbale ca modalitate de stabilire a unei relații între membrii unei perechi morfonologice.

3.3 Introducere

Generalizarea formalismelor bazate pe unificare este susținută de argumentele declarativității, reversibilității și reutilizabilității descrierii lingvistice și a generat, în timp, modele de limbă, bazate pe teoria structurilor de tip atribut-valoare.

Descrierile operate în interiorul morfologiei flexionare utilizează, de regulă, conceptul de paradigmă pentru a crea grupe de reguli ce caracterizează comportamentul flexionar al unei anumite submulțimi din vocabularul limbii respective. De exemplu, gramatica limbii române împarte verbele în grupe corespunzătoare conjugărilor I, a II-a, a III-a și a IV-a, deci regulile vizează particularitățile acestor tipuri de conjugări. Familiile formelor lexicale pe care cuvintele le pot avea sunt descrise prin combinații între forma bază sau rădăcină și paradigma căreia cuvântul îi aparține. Cuvintele neregulate al căror comportament iese din cadrul paradigmelor generale pot adeseori să fie descrise parțial prin referire la șabloane paradigmatic regulate, accesând lexiconul și regulile.

Structura unui cuvânt este similară comportamentului unui compozit ale cărui reguli de execuție sunt aplicate unei mulțimi de constante și variabile [27, 28]. Din perspectivă morfologică, acest compozit este o paradigmă și constă în inventarul

morfemelor și ansamblurilor de morfeme care alcătuiesc flexiunile, precum și în totalitatea schemelor formelor paradigmatică ale cuvintelor.

În lucrarea **A high-level morphological description language exploiting inflectional paradigms**, Anick și Artemieff prezintă o notație în care paradigma joacă un rol central- PDL (*Paradigm Description Language*) și care a fost utilizată ca bază pentru un analizor morfologic, o componentă a unui sistem multilingv de căutare a informației asistat de lexicon, AI-STARS.

PDL este alcătuit din trei componente principale: regulile de formare, ierarhia de moștenire a paradigmelor și regulile ortografice. În regulile de formare a cuvintelor se disting două categorii de șiruri lexicale:

- *rădăcinile*, respectiv orice formă care include baza lexicală primară a cuvântului;
- *afixele*, incluzând prefixe și sufixe ce se alătură unui cuvânt, modificându-i sensul într-un altul, de regulă, apropiat sau asociat.

Regulile de construcție a formelor lexicale noi sunt restricționate la cele cinci cazuri de mai jos:

- a. <formă lexicală> : <rădăcină> - <afix>
- b. <formă lexicală> : <rădăcină> + <afix>
- c. <formă lexicală> : + <afix> <rădăcină>
- d. <formă lexicală> : - <afix> <rădăcină>
- e. <formă lexicală> : <rădăcină>

Notația <formă lexicală> reprezintă șirul creat prin regula de derivare. Notația <rădăcină> reprezintă forma bază. Notația <afix> este un șir, prefix sau sufix, iar tipul său implică poziția înainte sau după <rădăcină> în regula lexicală. Operatorul (+ sau -) precede întotdeauna afixul.¹⁶ Forma lexicală rezultată poate fi folosită, la rândul ei, ca rădăcină în construcția altor cuvinte. În acest mod, crearea unei forme de suprafață poate fi descrisă prin intermediul unor operații de afixare sau de schimbare a rădăcinii, fiecare operație fiind descrisă prin câte o regulă.

¹⁶ Dacă este "+", atunci afixul este adăugat rădăcinii ca prefix sau sufix. Dacă este "-", afixul este îndepărtat de rădăcină.

O paradigmă, în PDL, este compusă dintr-o mulțime de reguli de formare lexicală care caracterizează, în mod colectiv, familia formelor de suprafață pentru acele cuvinte ce aparțin paradigmei. Pentru a captura similitudinile dintre paradigme și a evita redundanțele în descrierea limbajului, s-a introdus conceptul de *ierarhie a moștenirii paradigmelor*: dacă paradigma B se bazează pe paradigma A, atunci toate formele și regulile de formare definite pentru paradigma A se aplică și pentru paradigma B. De asemenea, B se diferențiază de A atunci când:

- se adaugă noi forme și reguli de construcție lexicală pentru forme care nu există în A;
- se suprascruie reguli de construcție lexicală pentru forme care există în A;
- o formă din A nu mai este aplicabilă în B, situație în care forma lexicală este ștearsă din B.

Concatenarea sau ștergerea unui afix poate rezulta, în multe cazuri, din schimbări ortografice (de exemplu, în engleză, multe lexeme terminate în vocală urmată de consoană vor dubla consoana finală atunci când este atașat un afix începând cu o vocală: *to clap- clapping*). În context PDL, fiecare regulă ortografică se aplică unui context ortografic specific rădăcinii și afixului pentru operațiile de concatenare și ștergere:

- reguli sufix, care descriu modificările ortografice aplicate la sfârșitul rădăcinii;
- reguli prefix, pentru a descrie modificările ortografice operate la începutul rădăcinii.

Organizarea comportamentului morfologic în termenii unor paradigme flexionare relaționate ierarhic ajută la caracterizarea explicită a similarităților și diferențelor dintre clasele de cuvinte.

Posibilitățile de implementare a analizei morfologice au în vedere următoarele aspecte [91]:

- Lexiconul conține toate formele unui cuvânt, analiza reducându-se la operația de căutare. Nu este o soluție viabilă din punctul de vedere al complexității;
- Folosirea unor automate și traductoare finite pentru reprezentarea lexiconului și a regulilor de formare¹⁷. Traductorul cu număr finit de stări este o generalizare a unui automat finit, în sensul că tranzițiile între stări nu se fac pe baza unui caracter, ci pe baza unor perechi de caractere. Aceste perechi de caractere fac parte dintr-un limbaj de intrare și unul de ieșire (sau, altfel spus, pe fiecare

¹⁷ vezi, de exemplu, XEROX: <http://www.xrce.xerox.com/research/mltt/>

tranziție, există o intrare și o ieșire), ceea ce deosebește funcționarea traductorului finit de cea specifică automatul finit care poate fi considerată o repetare a citirii unui simbol pe o bandă de intrare și trecerea într-o nouă stare.

- Lexiconul conține doar rădăcini și afixe. Regulile sunt folosite pentru identificarea formei cuvântului¹⁸. Analiza morfologică se face prin relaționarea a două niveluri de reprezentare a unui cuvânt: nivelul de suprafață (reprezentarea textuală) și nivelul de adâncime (reprezentarea de dicționar).

Morfologia pe două niveluri [52], [54] a înlesnit dezvoltarea descrierilor morfologice pentru limbile naturale, o dovadă semnificativă fiind sistemele *Kimmo standard* create pentru câteva limbi, incluzând finlandeza, suedeza, rusa, engleza și araba. Sistemele sunt alcătuite dintr-un lexicon în forma unei arborescențe de caractere (structura trie), având nodurile frunză adnotate cu informație morfologică, și un ansamblu de traductoare care codifică alternanțele morfologice. O formă lexicală flexionată e analizată prin punerea în corespondență a unei forme de intrare cu o secvență de forme lexicale, prin intermediul traductoarelor și prin compunerea unei ieșiri lexicale din adnotările conținute în nodurile frunză, aflate pe drumurile traversate în arbore.

Morfologia pe două niveluri are un caracter universal, deoarece modelează atât productivitatea inflexională, cât și cea de aglutinare; este o morfologie derivativă care tratează prefixe, rădăcini, sufixe. Pentru orice articol lexical, există o clasă de continuare, un spațiu în care prin model se restricționează natura elementelor lexicale ce pot apărea după el. Dicționarul devine un graf al claselor de continuare, unde fiecare clasă de continuare este un dicționar specializat (d.e., dicționarul sufixelor nominale). Din punct de vedere computațional, morfologia derivativă este foarte complexă, iar din punct de vedere lingvistic operează cu următoarele reguli cu două niveluri:

¹⁸ vezi, de exemplu, KIMMO: <http://www.sil.org/pckimmo/pc-kimmo.html>

- reguli de restricționare contextuală

$c_L:c_S \Rightarrow c_L:c_S' - c_L :c_S$ // simbolul c_L (caracter lexical) poate fi asociat cu simbolul c_S (caracter de suprafață), dacă este precedat de asociația $c_L:c_S'$ și succedat de asociația $c_L :c_S$

- reguli de restricționare a formei de suprafață

$c_L:c_S \Leftarrow c_L:c_S' - c_L :c_S$

- reguli compuse sau de echivalență, rezultate din combinația celor două reguli anterioare

$c_L:c_S \boxtimes c_L:c_S' - c_L :c_S$ // simbolul c_L (caracter lexical) obligatoriu se asociază cu simbolul c_S (caracter de suprafață), dacă este precedat de asociația $c_L:c_S'$ și succedat de asociația $c_L :c_S$

- reguli de excludere

$c_L:c_S / \Leftarrow c_L:c_S' - c_L :c_S$ // simbolul c_L (caracter lexical) nu poate fi asociat cu simbolul c_S (caracter de suprafață), dacă este precedat de asociația $c_L:c_S'$ și succedat de asociația $c_L :c_S$

Figura 2.1: Exemple de reguli tipice pentru morfologia pe două niveluri

- geminația (dublarea consoanei finale):

+: $CG \boxtimes CG - V:V$

refer + ing = referring

$\langle r:r \rangle \langle e:e \rangle \langle f:f \rangle \langle e:e \rangle \langle r:r \rangle \langle +: CG \rangle \langle i:i \rangle \langle n:n \rangle \langle g:g \rangle$,

unde: $CG = \{b, d, f, g, l, m, n, p, r, s, t\}$; $V = \{a, e, i, o, u\}$

- epenteza (inserția unei vocale la limita morfemului):

+: $e \boxtimes \{ \langle s:s \ c:c \rangle \ h:h \rangle \ s:s \ x:x \ z:z \} -s$

crash, branch + s = crashes, branches

bus + s = buses

box + s = boxes

buzz + s = buzzes

Aplicarea succesivă a regulilor nu este o abordare viabilă, pentru că ordinea este esențială. De aceea regulile se aplică în paralel, altfel spus, se face o combinație a tuturor regulilor, obținându-se un traductor între cele două niveluri.

Dacă modelul fono-morfologic pe două niveluri al lui Koskeniemi este un argument al formalizării procedurale operate în descrierile nivelurilor fonologice,

morfologice și ortografice, pe de altă parte, generalizarea formalismelor bazate pe unificare a condus la noi abordări ale analizei morfologice, bazate pe teoria structurilor de tip atribut-valoare.

Prima modelare și prima implementare din lume a morfologiei paradigmatică a unei limbi naturale, bazate pe unificare și structuri de tip atribut-valoare îi aparține profesorului Dan Tufiș, a cărui lucrare “It Would Be Much Easier if **Went** Were **Goed**” a primit cel mai mare scor la Conferința Europeană a Asociației de Lingvistică Computațională de la Manchester, în 1989.

Pe baza unui nucleu de cunoștințe morfo-lexicale, obținut cu ajutorul sistemului de învățare din exemple PARADIGM, în anul 1992 la Institutul de Cercetare pentru Inteligență Artificială s-a început o descriere exhaustivă a morfologiei limbii române, folosindu-se un limbaj descriptiv neutru și reversibil (FAVR- Flat Attribute Value Representation). Pentru validarea corectitudinii, acoperirii lingvistice, a reversibilității și a reutilizabilității descrierii paradigmatică pentru limba română, cercetătorii de la ICIA au optat pentru un formalism specific, respectiv ELU, dezvoltat la ISSCO- Geneva. Corecțiile și adăugirile făcute în mediul de execuție Mac-ELU, o versiune îmbunătățită a sistemului ELU, au fost operate și în FAVR, aceasta constituind probabil cea mai completă resursă lingvistică computațională a morfologiei limbii române [71, 72].

Paradigmele sunt clase de echivalență flexionară, în care cuvântul este privit ca rezultatul concatenării unei părți (relativ) invariante, numită temă, și a unei părți variabile, numită flexiune (terminație).

Structura generică a unui cuvânt în contextul mediului de dezvoltare Mac-Elu este descrisă de următoarea expresie regulată, în care sufixul vid este considerat un sufix gramatical propriu-zis: (prefix)* rădăcină (sufix-lexical)* sufix-gramatical.

Rădăcina împreună cu primul sufix lexical formează o temă lexicală, care- dacă e urmată de un alt sufix lexical, constituie o nouă temă, procesul de derivare continuând până când nu mai există niciun sufix lexical. Se numește temă lexicală implicită o rădăcină care nu este urmată de niciun sufix lexical.

Informația din intrarea lexicală corespunzătoare unei rădăcini are în vedere categoria sau categoriile gramaticale, forma leasă (asociată temei implicite), prefixe acceptabile și modul de sufixare adiacentă (lexicală sau gramaticală). De asemenea,

temele implicite au asociată informație sintactică (de exemplu, liste de subcategorizare) și semantică (structuri de theta-roluri, restricții selecționale).

Paradigmele gramaticale ale verbului sunt în număr de 48, 3 dintre acestea fiind specifice verbelor auxiliare, și reprezintă combinații ale paradigmelor parțiale corespunzătoare modurilor și timpurilor simple existente în limba română (în număr de 107 paradigme parțiale). Pe lângă acestea există un număr de 27 de paradigme lexicale care atașează temelor verbale sufixe lexicale care schimbă categoria gramaticală a verbului obținându-se substantive, adjective și adverbe. În figurile 1 și 2 sunt prezentate în limbajul ELU o paradigmă (gramaticală) totală pentru teme verbale și respectiv o paradigmă parțială a acesteia (una din cele șase paradigme parțiale corespunzătoare indicativului mai-mult-ca-perfect):

```
# paradigm verb1
!Verb !type(main)
$indic_prez_1
$indic_imperf_1
$indic_perfsim_1
$indic_mmcperf_1
$conj_prez_1
$imper_prez_1
$infin_prez_1
$part_1
$gerund_2
```

```
# paradigm indic_mmcperf_1
!VTensed(past-perfect,indicative)
    asem v {+past} !my_Vagr(singular,1,_)
    aseși v {+past} !my_Vagr(singular,2,_)
    ase v {+past} !my_Vagr(singular,3,_)
    aserăm v {+past} !my_Vagr(plural,1,_)
    aserăți v {+past} !my_Vagr(plural,2,_)
    aseră v {+past} !my_Vagr(plural,3,_)
```

Figura 2.2: Paradigmă verbală globală Figura 2.3: Paradigmă verbală parțială

Declinarea substantivului comun a fost modelată prin intermediul a 52 de paradigme flexionare totale, 14 corespunzătoare genului masculin, 22 corespunzătoare genului feminin și 16 genului neutru. Prin intermediul restricțiilor, aceste paradigme pot fi folosite și pentru substantivele proprii, existând și un număr de 7 paradigme specifice substantivelor proprii. La rândul lor, aceste paradigme totale sunt combinațiile unui număr de 78 paradigme parțiale corespunzătoare genurilor, numerelor și tipului declinării (articulată și respectiv, nearticulată). Pe lângă acestea, există clase de continuare specifice sufixelor diminutive (47), moționale (11), augmentative (7), acestea neschimbând categoria gramaticală a substantivului, și alte sufixe lexicale (46).

Pentru celelalte categorii gramaticale flexionare (articol, pronume și adjectiv nehotărât, pronume și adjectiv demonstrativ), au fost scrise 19 paradigme flexionare, iar declinarea adjectivelor și numeralelor care, prin conversiune, devin substantive sau adjective, se face după paradigmele flexionare ale numelor comune.

3.4 Paradigmele nominale în limba română

Deflexionarea este procesul în care se stabilește forma bază a cuvântului de analizat și codul morfologic al acestuia. Pentru limba română, deflexionarea este împărțită în două submodule, deflexionarea nominală și deflexionarea verbală, realizate diferit și din punctul de vedere al problematicii, și al rezolvării. Ce au totuși în comun aceste două submodule este faptul că ambele folosesc terminațiile cuvintelor pentru a încadra cuvintele în anumite grupe, de unde, în urma validării formei bază le atașează anumite informații morfologice.

Obținerea paradigmatelor nominale pentru limba română, prezentată în această teză, este rezultatul unui proces științific dezvoltat pe parcursul pregătirii lucrării în mai multe etape:

- *Etapa de cercetare*, în care am urmărit obținerea tuturor formelor flexionare pentru un număr de aproximativ 2000 de substantive de toate genurile, parcurgând traseul morfologic: ((G.D, plural, articulat) → (N.Ac., plural, articulat) → (N.Ac., plural, nearticulat) → (G.D., singular, articulat) → (N.Ac., singular, articulat) → (N.Ac., singular, nearticulat)).
- *Etapa de analiză*, în care am propus două metode pentru deflexionare, prima fiind orientată spre acoperire lingvistică, a doua spre optimizarea timpului de obținere a lemei. Tot în această etapă am identificat și grupat anumite terminații specifice cazurilor corespunzătoare pașilor traseului morfologic prezentat mai sus. Ele vor fi necesare mai târziu când, după validarea formei bază, cuvântul va primi un cod morfologic cu date privitoare la număr, gen, caz și articol.

- *Etapa de evaluare* a paradigmelor create, în care am considerat aproximativ 1000 de substantive în toate formele flexionare, am identificat paradigma lor proprie și am parcurs ordinea operațiilor indicate pe liniile acestora.
- *Includerea paradigmelor într-o aplicație* pentru verificarea finală, realizând un program în mediul Delphi de traducere automată, pe direcția română-engleză.

3.4.1 Prima metodă de deflexionare nominală

Scopul primei metode create pentru deflexionarea nominală a fost acoperirea lexicală. Am procedat astfel la clasificarea terminațiilor, am numit restricțiile de Nominativ-Acuzativ, plural, nearticulat ca fiind identificatorii paradigmelor și am impus operații de tip *scad șir*, *adun șir*, *alternanțe fonetice* în interiorul regulilor lexicale și al relațiilor morfologice. Rezultatul a fost crearea unor paradigme nominale ce folosesc o serie de grupe de terminații, alta de restricții și operația de unificare de șiruri pentru analiza morfologică.

Paradigma substantivului este exprimată prin relațiile morfologice și regulile lexicale de tipul *substantiv [Tipul, Șir: Proprietăți, RL, S]*, unde Proprietățile pot include ecuații de șir, iar RL și S sunt liste ale numelor regulilor lexicale și, respectiv, ale specificațiilor de șir. Ideea care stă la baza creării paradigmelor este de a îngloba operații de tipul celor descrise mai sus, într-o anumită ordine, fiecare specifică unei terminații, astfel încât, indiferent de terminația identificată, asupra cuvântului să se efectueze în mod automat un set bine stabilit de operații.

În figura 2.4 este descrisă o paradigmă de flexionare pentru substantivele masculine cu terminațiile {și, ști} pentru formele de Nominativ-Acuzativ, plural, nearticulat, iar tabela 2.1 particularizează această paradigmă pentru cuvintele *urs* și *artilierist*.

Prima linie completează tipul substantivului, categoria genului și precizează terminația de restricție. A doua linie ordonează regulile lexicale prin care se vor genera formele flexionate, iar ultima pornește de la terminația de G.D. plural, articulat, continuă cu operațiile de *scad șir*, *adun șir*, *alternanțe fonetice* până la ultimul câmp completat:

Figura 2.4: Paradigmă nominală pentru substantive masculine cu terminațiile {*și, ști*}
 ([nume comun, masculin, restricție NAc pl neart #și; #ști],
 [GD pl art NAc pl art NAc pl neart GD sg art NAc sg art NAc sg neart]
 [#lor -lor, +i -i -i, ș->s, +ului -ui -ul])

Tabela 2.1: Comportamentul flexionar al substantivelor *urs* și *artilerist*

Informația morfologică	Operația de șir			Terminația	Cuvântul-exemplu
	Scad șir	Alternanță	Adun șir		
GD pl art	-	-	-	lor	urșilor artileriştilor
NAc pl art	lor	-	i	ii	urșii artileriştii
NAc pl neart	i	-	-	și ști	urși artilerişti
GD sg art	i	ș->s	ului	ului	ursului artileristului
NAc sg art	ui	-	-	ul	ursul artileristul
NAc sg neart	ul	-	-	#	urs artilerist

Conform principiului de funcționare a paradigmelor prezentat mai sus, oricare ar fi terminația identificată din coloana de terminații, asupra cuvântului se vor efectua toate operațiile prezentate în tabel, de la acea terminație până la sfârșitul tabelului.

Regula de aplicare a alternanțelor este de a se căuta șirul ce trebuie înlocuit de la dreapta la stânga, iar înlocuirea se efectuează o singură dată.

Relația de subsumare pe care se bazează descrierea paradigmelor definește o ordine parțială asupra specificațiilor de șir, astfel încât o specificație S subsumează o alta S' dacă toate instanțele bază din S' sunt, de asemenea, instanțe în S. Exemplele de mai sus validează paradigmele nominale nu doar pentru substantivele ce figurează ca intrare

lexicală (*urs, artilerist*), ci grupează toate numele comune având aceeași terminație de restricție (de exemplu, *urs, cais, pas; artilerist, jurist, genist*).

În Figura 2.5 voi descrie o paradigmă flexionară pentru substantivele feminine cu terminațiile de N.Ac. plural nearticulat: *ete, ețe, epe, erze, ese, etre, este*, iar în Tabela 2.2 voi particulariza această paradigmă prin substantive ce îi respectă restricția:

Figura 2.5: Paradigmă flexionară pentru substantivele feminine cu terminațiile *ete, ețe, epe, erze, etre, este*

[nume comun, feminin, restricție NAc pl neart #ete; #ețe; #epe; #erze; #etre; #este],

[GD pl art NAc pl art NAc pl neart GD sg art NAc sg art NAc sg neart]
[#lor -lor, +le -le +i -ei, e->a, +a -a, +ă]

Tabela 2.2: Comportamentul flexionar al substantivelor ce respectă restricția *ete, ețe, epe, erze, etre, este*

Informația morfologică	Operația de șir			Terminația	Cuvântul-exemplu
	Scad șir	Alternanță	Adun șir		
GD pl art	-	-	-	lor	fetelor piețelor iepelor berzelor pietrelor nevestelor
NAc pl art	lor	-	le	le	fetele piețele iepele berzele pietrele nevestele
NAc pl neart	le	-	-	ete ețe epe erze etre este	fete piețe iepe berze pietre neveste
GD sg art	-	-	i	ei	fetei pieței iepei berzei pietrei nevestei

NAc sg art	ei	e->a	a	a	fata piața iapa barza piatra nevasta
NAc sg neart	a	-	ă		fată piață iapă barză piatră nevastă

Figura 2.6: Pașii de analiză a formei *fetele*

Intrare: *fetele*; terminație recunoscută: *le*; informație morfologică: NAc pl art

<scad șir>: [-le] fete

<verific restricția>: citesc subșirul [ete]

<adun șir>: [+i] fetei

<scad șir>: [-ei] fet

<adun șir>: [+a] feta

<alternanțe fonetice>: [e->a] fata

<scad șir>: [-a] fat

<adun șir>: [+ă] fată

Ieșire: fată; forma se găsește în dicționar => se validează *ete* ca restricție pentru substantiv comun feminin de forma *fetele*; se validează *fată* ca lemă a cuvântului de intrare, acesta primind și informație morfologică.

Restricția-nucleu este activată atunci când este în întregime conținută în șirul de intrare. În exemplul *zilele*, este activat nucleul *le*, pe care programul îl găsește în baza de date ca restricție. Lexemul de intrare *fetele*, având aceeași terminație, dar analizat conform pașilor paradigmei corespunzătoare restricției *le*, va genera o formă nevalidă:

Figura 2.7: Identificarea șirului *le* ca restricție pentru forma *zilele*

Intrare: *zilele*; terminație recunoscută: *le*; informație morfologică: NAc pl art

<scad șir>: [-le] zile

<verific restricția>: citesc subșirul [le]

<adun șir>: [+i] zilei
 <scad șir>: [-lei] zi
 <adun șir>: [+ua] ziua
 <scad șir>: [-ua] zi

Leșire: zi; se validează *le* ca restricție pentru nume comun feminin de forma *zilele*

Figura 2.8: Activarea restricției *le* pentru forma *fetele*; fără validarea formei finale

Intrare: fetele

<verific restricția>: citesc subșirul [le]
 <adun șir>: [+i] *fetelei
 <scad șir>: [-lei] *fete
 <adun șir>: [+ua] *feteua
 <scad șir>: [-ua] *fete

Leșire: fete; nu se validează *le* ca restricție pentru nume comun feminin de forma *fetele*

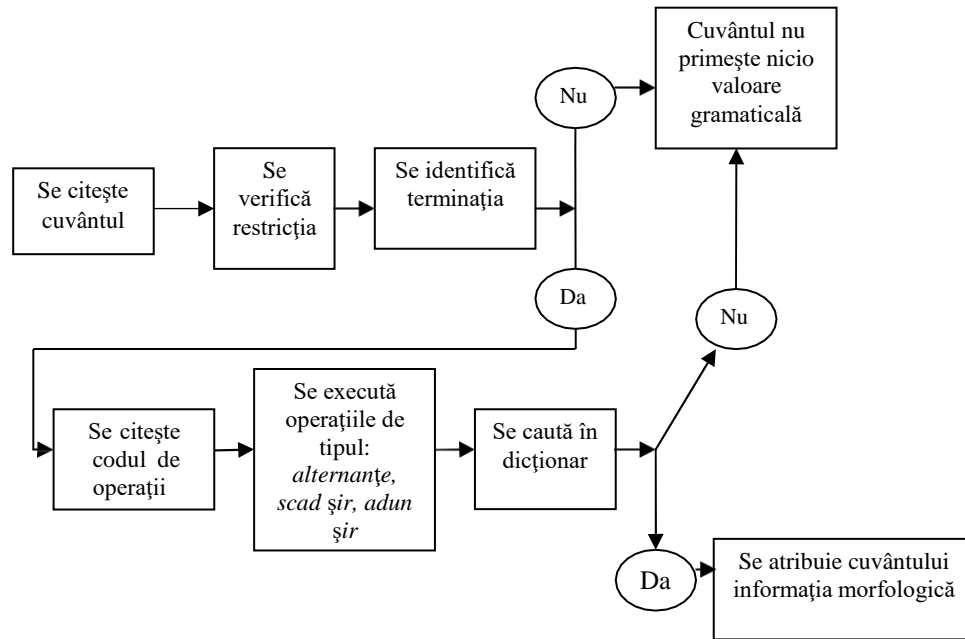
3.4.2 Algoritm de deflexionare nominală

Se consideră cuvântul de intrare, introdus de utilizator, se citește terminația și se compară cu unul din șirurile ce clasifică restricțiile de N-Ac., plural, nearticulat. În cazul de față, terminațiile reprezintă șiruri de caractere, de diferite mărimi, în funcție de lungimea cuvântului, formate crescător, începând cu ultima literă, la care se adaugă, pe rând, următoarea literă, în sensul de la dreapta la stânga. Acestea se separă de la începutul procedurii și se salvează într-o listă de terminații. Se verifică, pe rând, existența acestor terminații în paradigmele de deflexionare, începând cu terminația cea mai mare.

Dacă aceasta este găsită, algoritmul citește din liniile paradigmei prima operație completată, fie *adun șir*, fie *scad șir*, fie *alternanțe fonetice*. Se efectuează apoi restul de operații, pe rând, în ordinea în care au fost introduse în paradigmă, după care, în final, se obține o formă lexicală considerată o potențială intrare de dicționar. Se verifică din nou în dicționar existența formei lexicale obținute și, dacă aceasta este găsită, atunci cuvântul flexionat primește valorile asociate câmpului morfologic de la care s-a pornit analiza.

Dacă nu se găsește forma bază, se trece la o altă paradigmă care acceptă una din terminațiile cuvântului de intrare. Analiza se oprește în momentul în care substantivul introdus de utilizator primește informație de flexiune nominală.

Procedura de deflexionare prezintă următorii pași (Figura 2.9):



Algoritmul de deflexionare bazat pe paradigme nominale este identic pentru toate tipurile de substantive comune: masculine, feminine, neutre. Aceiași pași de analiză sunt urmăriți și pentru obținerea formei bază în cazul adjectivelor calificative.

Deflexionarea nominală folosește ca terminații șirurile de 1 până la 4 caractere, în funcție de cuvânt, considerate de la dreapta la stânga.

În contextul mediului de programare Delphi, am realizat reprezentarea pentru o restricție creând un obiect denumit “Regulă”, ce conține două proprietăți (o listă de identificatori și o listă de operații) și trei metode (două pentru adăugarea în liste și una pentru instanțierea propriu-zisă):

```

type TRegula = class
  listaOperatii : TList;
  listaIdentif : TList;

```

```

public
  constructor Create;
  procedure adaugaOperatie(operatie: TOperatie);
  procedure adaugaIdentificator(identif : TIdentificator);
  procedure proceseaza(var s,d:string) ;
end;

```

Identificatorii care se adaugă în listaIdentif, precum și operațiile care se adaugă în listaOperatii sunt definiți ca obiecte separate:

```

type TIdentificator = class
  public
    legatura1, legatura2, terminatie: string;
    Constructor Create(leg1, leg2, term: string);
end;

```

unde *legatura1* reprezintă codul cu care se va identifica o anumă operație din lista de operații, *legatura2* reprezintă valoarea morfologică pe care o va primi cuvântul, iar *term* reprezintă șirul cu care se va compara terminația curentă din lista de terminații;

```

type TOperatie = class
  public
    legatura1, terminatie: string;
    Constructor Create(leg1, term: string);
    procedure modifica(var s, term:String); virtual ; abstract;
end;

```

unde *legatura1* reprezintă codul de recunoaștere a operației în funcție de identificator, *terminatie* înseamnă valoarea terminației curente, iar *procedure modifica* este metoda acestui obiect ce conduce prin definirea sa la ștergerea unui subșir (operația *sterge_terminatie*) sau adăugarea unui șir (operația *adaug_terminatie*) la sfârșitul șirului de intrare, precum și la înlocuirea unui subșir cu altul (operația *alternanta*).

2.2.3. A doua metodă de deflexionare nominală

A doua metodă de punere în aplicație a morfologiei paradigmatică este realizată pe baza aceluiași reguli de deflexionare create în etapa de cercetare. Diferența constă într-o altă abordare a paradigmelor și o optimizare a operațiilor necesare.

În cazul de față se renunță la obiectele de tip *Regulă* ce înglobau atât serii de identificatori, cât și serii de operații, utilizându-se un nou obiect de tip *Restricția*, ce conține doar restricția prin care se identifică o paradigmă și alternanța specifică acesteia (Figura 2.10). Procedura *Substantiv* se păstrează, dar nu ca metodă a unui obiect, ci ca o procedură de sine stătătoare.

```
type TRestricția=class
    public
    restrictie, cod, alt11, alt12, alt21, alt22:string;
    constructor create (r,c,a11,a12,a21,a22:string);
```

Figura 2.10: Definirea obiectului *Restricția*

Principiul de realizare a procedurii de deflexionare nominală se focalizează asupra terminațiilor cu informație morfologică, în funcție de care se impun anumite modificări ale șirului de intrare. Aceste terminații sunt particulare cazurilor: GD plural <lor>; NAc plural, articulat <ii, le>; GD singular <ului, lui; ei, ii>; NAc singular, articulat <le, ul, l; a, ua>.

Ca și în metoda prezentată anterior, restricțiile sunt identificatori ai paradigmelor și constituie subșiruri ale formelor de NAc plural, nearticulat. În acest caz ele sunt reprezentate în aplicație prin obiectele de tip *Restricție* și sunt adunate într-o listă *lrestricții: TList*.

În funcție de fiecare terminație, șirul de intrare este normalizat prin diferite seturi de operații la forma de NAc plural, nearticulat. În acest moment se deschide lista de restricții și se verifică dacă vreunul din subșirurile ce reprezintă restricțiile este inclus în șirul corespunzător formei de NAc plural, nearticulat. Dacă se găsesc unul sau mai multe subșiruri, conform fiecărei restricții, lexemul este supus anumitor operații pentru a fi adus la forma lemă care, la rândul ei, este salvată într-o listă temporară.

La finalizarea acestor pași, rezultatele obținute se caută în baza de date pentru validare. În urma cercetării, am observat că, la verificarea restricțiilor, pot fi identificate maximum două restricții pentru o formă, iar, din acestea, în urma căutării în baza de date, se validează un singur rezultat (Figura 2.11):

Figura 2.11: Situație conflictuală cu două restricții potențiale pentru un cuvânt
forma de intrare *tașilor, soldașilor*

- se identifică terminația *lor*;
- se scade terminația, obținându-se, conform tabelului de paradigme, forma de NAc, plural, nearticulat: *tași, soldași*;
- se verifică restricțiile => *ași, și*;
- se realizează alternanțele fonetice:
 - ași*: alternanța *și* -> *tă* => *tată*, **soldată*
 - și*: alternanța *și* -> *t* => **tat*, *soldat*
- se caută rezultatele în baza de date și se validează forma *tată*, cu traducerea *father*, respectiv *soldat*.

Această metodă reduce timpul de prelucrare a paradigmele prin reducerea numărului de operații aplicate șirului de intrare, precum și prin modul de selecție a setului de operații ce trebuie aplicate (Figura 2.12):

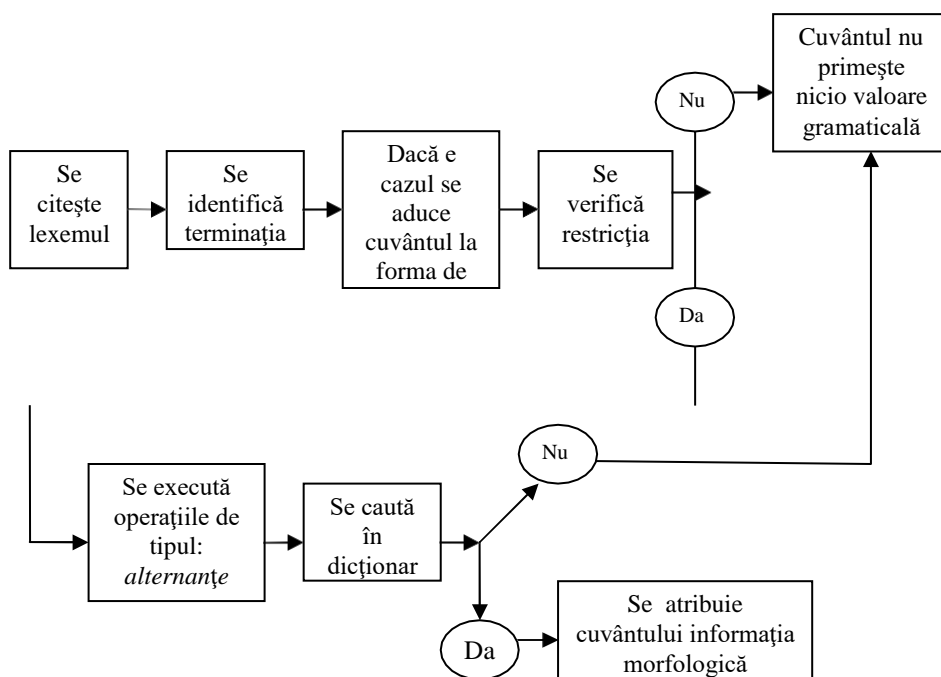


Figura 2.12: Etapele procesului de deflexionare nominală. A doua metodă

2.3. Paradigmele verbale în limba română

2.3.1 Prezentarea cadrului de lucru

Deflexionarea verbală este un proces mai complex decât cel pentru deflexionarea nominală prin numărul de paradigme existent și prin numărul de condiții ce trebuie îndeplinite în același timp pentru validarea formei bază. Vom prezenta în această lucrare paradigmele verbale create pentru indicativ, conjugarea I, a II-a, a III-a, respectiv a IV-a. Ca și în cazul paradigmelor nominale, Proprietățile includ ecuații de șir, iar regulile lexicale și specificațiile de șir descriu formele verbale tranzitive de la verbul de intrare, conjugat, la cel existent în dicționar la infinitiv, formă scurtă.

Pentru a optimiza timpul de obținere a formelor finale relației verb conjugat- verb de dicționar, am organizat terminațiile în 25 de grupe distincte după comportamentul morfonologic pentru persoană, număr și timp.

Pentru verbele de conjugarea I există 6 grupe de verbe, dintre care ultima se referă la verbul „a da”, tratat în regim de verb neregulat. Ținând cont de faptul că terminațiile de prezent sunt cele care deosebesc verbele de conjugarea I, considerate verbe regulate în cadrul acestei analize, am creat o clasă generală de terminații, având exclusiv terminații verbale pentru timpurile simple ale modului indicativ: imperfect, perfect simplu și mai-mult-ca-perfect (Tabela 2.3):

Persoana	Numărul	Terminații Imperfect	Terminații Perfect simplu	Terminații Mai-mult-ca- perfect
persoana I	singular	am	ai	asem
persoana a II-a		ai	ași	aseși
persoana a III-a		a	ă	ase
persoana I	plural	am	arăm	aserăm
persoana a II-a		ați	arăți	aserăți
persoana a III-a		au	ară	aseră

Tabela 2.3: Clasă generală de terminații pentru verbele de conjugarea I (ConjI.CGT)

Clasa generală de terminații este completată prin 4 clase parțiale de terminații (CPT), create pentru formele temporale de prezent (Tabelul 2.4):

	Persoana și numărul	Terminații Prezent			
		ConjI.CPT1	ConjI.CPT2	ConjI.CPT3	ConjI.CPT4
Verb de conjugarea I	pers. I, nr. sg.	#	ez	i	ez
	persoana a II-a, nr. sg.	i	ezi	i	ezi
	persoana a III-a, nr. sg.	ă	ează	ă	ază
	pers. I, nr. pl.	ăm	ăm	ăm	em
	persoana a II-a, nr. pl.	ați	ați	ați	ați
	persoana a III-a, nr. pl.	ă	ează	ă	ază

Tabela 2.4: Clase parțiale de terminații

Prima grupă de verbe de conjugarea I folosesc terminațiile din clasa generală, completate cu prima clasa parțială de terminații, pentru obținerea formelor conjugate de indicativ (de exemplu, verbele *a striga*, *a ridica*).

A doua grupă de verbe de conjugarea I diferă de comportamentul morfologic al clasei precedente prin prezența alternanțelor fonetice, prezentate în liniile tabelului de mai jos:

	Identificator de paradigmă	Paradigmă verbală: ConjI.CGT + ConjI.CPT 1 + Alternanțe fonetice	Exemplu
		Alternanțe fonetice	
Verbe de conjugarea I, a doua grupă (ConjI.V2)	ăut	<i>a</i> -> <i>ă</i> prezent, persoanele 1, 2, 3, 6	a căuta
		<i>ț</i> -> <i>t</i> prezent, persoana 2	
	ăt	<i>ț</i> -> <i>t</i> prezent, persoana 2	a arăta
		<i>a</i> -> <i>ă</i> prezent, persoanele 3, 6	
	ânc	<i>ă</i> -> <i>â</i> prezent, persoanele 1, 2, 3, 6	a mânca
		<i>nân</i> -> <i>n</i> prezent, persoanele 1, 2, 3, 6	
	tin	<i>a</i> -> <i>ă</i> prezent, persoanele 1, 2, 3, 6	a clătina
	ăsc	<i>a</i> -> <i>ă</i> prezent, persoanele 1, 2, 3, 6	a căsca
		<i>șt</i> -> <i>sc</i> prezent, persoana 2	
	ătr	<i>a</i> -> <i>ă</i> prezent, persoanele 1, 2, 3, 6	a lătra
	ent	<i>ț</i> -> <i>t</i> prezent, persoana 2	a prezenta
		<i>i</i> -> <i>e</i> prezent, persoanele 1, 2, 3, 6	
lt, et, ânt, ut, it	<i>ț</i> -> <i>t</i> prezent, persoana 2	a asculta a cugeta a cânta a muta	

			a merita
erg, emn, er, est, ers, ec, eg, eț	<i>ea</i> -> <i>e</i> prezent, persoanele 3, 6		a alerga a pleca a lega a îngheța
uc	<i>o</i> -> <i>u</i> prezent, persoanele 1, 2, 3, 6		a juca
ăl	<i>a</i> -> <i>ă</i> prezent, persoanele 3, 6		a spăla
	<i>e</i> -> <i>ă</i> prezent, persoana 2		
pt	<i>ea</i> -> <i>e</i> prezent, persoanele 3, 6		a aștepta
	<i>ț</i> -> <i>t</i> prezent, persoana 2		
ug	<i>o</i> -> <i>u</i> prezent, persoanele 1, 2 <i>oa</i> -> <i>u</i> prezent, persoanele 3, 6		a ruga

Tabela 2.5: Alternanțele fonetice aplicate celei de a doua grupe de verbe de conjugarea I

A treia grupă de verbe combină clasa generală de terminații (ConjI.CGT) și a doua clasă parțială de terminații (ConjI.CPT 2). Această grupă nu se combină cu linii de alternanțe fonetice în scopul conjugării, exceptând verbul *a naviga*, pentru care există alternanța fonetică *gh* -> *g* (Tabela 2.6):

	Identificator de paradigmă	Exemplu
Verbe de conjugarea I, a treia grupă (ConjI.V3)	on, or, ot	a colabora a dona a nota
	nt, u, r, nț	a evalua a contura a înștiința
	ul, rm, dr	a calcula a urma a încadra
	aj, am	a încuraja a programa
	id, iz, in	a candida a sensibiliza a examina
	es, est, ers	a interesa a manifesta a traversa
	cet, ct	a cerceta a proiecta

Tabela 2.6: A treia grupă de verbe de conjugarea I (ConjI.V3)

A patra grupă de verbe de conjugarea I combină clasa generală de terminații și a treia clasă parțială de terminații, fără să aibă alternanțe fonetice. Grupa este specifică pentru identificatorul de paradigmă *nu* (d.e., *a continua*).

Ultima grupă de verbe regulate de conjugarea I sunt analizate după paradigma {ConjI.CGT + ConjI.CPT 4¹⁹}, fără alternanțe fonetice. Grupa aceasta particularizează verbele cu identificator de paradigmă *i* (d.e., *a negocia*).

În continuare, am împărțit verbele de conjugarea a II-a în 4 grupe de verbe regulate și 2 grupe de verbe neregulate, create pentru verbele *a vrea*, *a avea*. Prima grupă de verbe se caracterizează prin următoarele terminații pentru timpurile simple ale modului indicativ (prezent, imperfect, perfect simplu, mai-mult-ca-perfect):

Verbe de conjugarea a II-a, prima grupă (ConjII.V1)	Persoana și numărul	Term. Prezent	Term. Imperfect	Term. Perfect simplu	Term. Mai-mult-ca-perfect
	pers. I, nr. sg.	#	eam	ui	usem
	persoana a II-a, nr. sg.	i	eai	uși	useși
	persoana a III-a, nr. sg.	e	ea	u	use
	pers. I, nr. pl.	em	eam	urăm	userăm
	persoana a II-a, nr. pl.	eți	eați	urăți	userăți
	persoana a III-a, nr. pl.	#	eau	ură	useră

Tabela 2.7: Terminații pentru prima grupă de verbe de conjugarea a II-a

La nivelul acestei prime grupe de verbe se diferențiază 5 subgrupe, în funcție de liniile de alternanțe implicate în conjugare:

subgrupa 1: {
 identificator: *ăr, ăc*;
 alternanță fonetică: *a* -> *ă* prezent, persoanele 1, 2, 3, 6;
 exemplu: *a părea*, *a tăcea*
 }

subgrupa 2: {
 identificator: *ed*;
 alternanțe fonetice: *ă* -> *e* prezent, persoanele 1, 6; perfect simplu și mai-mult-ca-perfect, persoanele 1, 2, 3, 4, 5, 6;
z -> *d* prezent, persoana 2; perfect simplu și mai-mult-ca-perfect, persoanele 1, 2, 3, 4, 5, 6;
 exemplu: *a vedea*

¹⁹ Clasa generală de terminații și a patra clasă parțială de terminații

- subgrupa 3: {
 identificator: *şed*;
 alternanțe fonetice: *z* -> *d* prezent, persoana 2; perfect simplu și mai-mult-ca-perfect,
 persoanele 1, 2, 3, 4, 5, 6;
 exemplu: a şede
 }
- subgrupa 4: {
 identificator: *ăd*;
 alternanțe fonetice: *a* -> *ă* prezent, persoanele 1, 2, 3, 6;
z -> *d* prezent, persoana 2; perfect simplu și mai-mult-ca-perfect, persoanele 1, 2,
 3, 4, 5, 6;
 exemplu: a cădea
 }
- subgrupa 5: {
 identificator: *ut*;
 alternanțe fonetice: *ț*-> *t* prezent, persoana 2;
o -> *u* prezent, persoanele 1, 2, 6; *oa* -> *u* prezent, persoana 3;
 exemplu: a putea }

Pentru următoarele grupe de verbe regulate de conjugarea a II-a, respectiv 2, 3 și 4, nu există alternanțe fonetice. Ele implică existența identificatorilor de paradigmă *gh* (ConjII. V2), *cr* (ConjII. V3), *b* (ConjII. V4):

Verbe de conjugarea a II-a, a doua grupă (ConjII. V2)	Persoana și numărul	Term. Prezent	Term. Imperfect	Term. Perfect simplu	Term. Mai-mult-ca-perfect
	pers. I, nr. sg.	ez	eam	eai	easem
	persoana a II-a, nr. sg.	ezi	eai	eași	easeși
	persoana a III-a, nr. sg.	ează	ea	e	ease
	pers. I, nr. pl.	em	eam	earăm	easerăm
	persoana a II-a, nr. pl.	eți	eați	earăți	easerăți
	persoana a III-a, nr. pl.	ează	eau	eară	easeră

Verbe de conjugarea a II-a, a treia	Persoana și numărul	Term. Prezent	Term. Imperfect	Term. Perfect simplu	Term. Mai-mult-ca-perfect
	pers. I, nr. sg.	eez	eam	eai	easem
	persoana a II-a, nr. sg.	eezi	eai	eași	easeși

grupă (ConjII.V3)	persoana a III-a, nr. sg.	eează	ea	e	ease
	pers. I, nr. pl.	eăm	eam	earăm	easerăm
	persoana a II-a, nr. pl.	eați	eați	earăți	easerăți
	persoana a III-a, nr. pl.	eează	eau	eară	easeră

Verbe de conjugarea a II-a, a patra grupă (ConjII.V4)	Persoana și numărul	Term. Prezent	Term. Imperfect	Term. Perfect simplu	Term. Mai-mult- ca-perfect
	pers. I, nr. sg.	eau	eam	ăui	ăusem
	persoana a II-a, nr. sg.	ei	eai	ăuși	ăuseși
	persoana a III-a, nr. sg.	ea	ea	ău	ăuse
	pers. I, nr. pl.	em	eam	ăurăm	ăuserăm
	persoana a II-a, nr. pl.	eți	eați	ăurăți	ăuserăți
	persoana a III-a, nr. pl.	eau	eau	ăură	ăuseră

Tabela 2.8: Terminațiile pentru grupele de verbe 2-4, conjugarea a II-a (ConjII.V2, ConjII.V3, ConjII.V4)

În analiza verbelor de conjugarea a III-a am identificat patru grupe de verbe regulate și am construit patru clase parțiale de terminații, ținând cont de terminațiile identice ale verbelor pentru timpurile prezent și imperfect (ConjIII.CPT 1), perfect simplu și mai-mult-ca-perfect (ConjIII.CPT 2, ConjIII.CPT 3) și pentru timpul prezent (ConjIII.CPT 4).

Prima clasă parțială de terminații este specifică verbelor din grupele ConjIII.V1 și ConjIII.V2, care au identificatorii și terminațiile verbale de prezent și imperfect din tabela 2.9:

Clasă parțială de terminații 1 (ConjIII.CPT 1)	Identificatori de paradigmă	Terminații Prezent		Terminații Imperfect	
	ConjIII. V1: { ag, arg // exemplu: a trage, a sparge; it // exemplu: a trimite; oac, oarc, oat // exemplu: a coace; a toarce, a scoate; ân // exemplu: a rămâne; un // exemplu: a pune; up // exemplu: a întrerupe; }	#	// pers. 1	eam	// pers.
	i	// pers. 2	1		
	e	// pers. 3	eai	// pers.	
	em	// pers. 4	2		
	eți	// pers. 5	ea	// pers.	
	#	// pers. 6	3		
ConjIII. V2: { rn, rd // exemplu: a cerne, a pierde; ac, at // exemplu: a face, a bate; er, ec, ern, ep // exemplu: a începe, a cere, a petrece; așt, eșt // exemplu: a naște, a crește; oas, oașt // exemplu: a coase, a cunoaște; in // exemplu: a obține }			eam	// pers.	
			4		
			eați	// pers.	
			5		
			eau	// pers.	
			6		

Tabela 2.9: Clasă parțială de terminații

Clasele parțiale de terminații ConjIII.CPT2 și ConjIII.CPT3 cuprind terminațiile pentru timpurile perfect simplu și mai-mult-ca-perfect ale grupelor verbale ConjIII.V1 și ConjIII.V3 respectiv ConjIII.V2 și ConjIII.V4, iar clasa parțială de terminații ConjIII.CPT4 include terminațiile pentru timpul prezent ale verbelor din ConjIII.V4 (Tabela 2.10):

Clasă parțială de terminații 2 (ConjIII.CPT 2)	Identificatori de paradigmă	Terminații Perfect simplu		Terminații Mai-mult-ca-perfect	
	ConjIII. V1: { identificatorii din tabelul anterior }	sei	// pers.	sesem	// pers.
	1		1		
	seși	// pers.	seseși	// pers.	
	2		2		
	se	// pers. 3	sese	// pers.	
	serăm	// pers. 4	3		
	serăți	// pers. 5	seserăm	// pers.	
	seră	// pers. 6	4		
			seserăți	// pers.	
			5		
			seseră	// pers.	
			6		

Clasă	Identificatori de paradigmă	Terminații Perfect simplu	Terminații Mai-mult-ca-perfect
Clasă parțială de terminații 3 (ConjIII.CPT 3)	ConjIII. V2: { identificatorii din tabelul anterior }	ui // pers. 1 uși // pers. 2 u // pers. 3 urăm // pers. 4 urăți // pers. 5 ură // pers. 6	usem // pers. 1 useși // pers. 2 use // pers. 3 userăm // pers. 4 userăți // pers. 5 useră // pers. 6
	ConjIII. V4: { pl // exemplu: a umple }		

Clasă	Identificatori de paradigmă	Terminații Prezent
Clasă parțială de terminații 4 (ConjIII.CPT 4)	ConjIII. V4: { pl // exemplu: a umple }	u // pers. 1 v // pers. 2 e // pers. 3 em // pers. 4 eți // pers. 5 u // pers. 6

Tabela 2.10: Clasele parțiale de terminații ConjIII.CPT 2, ConjIII.CPT 3, ConjIII.CPT 4

Pentru primele două grupe de verbe regulate am identificat liniile de alternanțe fonetice de mai jos, iar grupele 3 și 4 se conjugă fără alternanțe fonetice:

grupa 1: {
subgrupa 1: {
identificator: *ag, arg*;
alternanțe fonetice: *ă -> a* imperfect, perfect simplu, mai-mult-ca-perfect, persoanele 1-6;
 # -> *g* perfect simplu, mai-mult-ca-perfect, persoanele 1-6
 }
subgrupa 2: {
identificator: *it*;
alternanțe fonetice: *ț -> t* prezent, persoana 2;
 # -> *t* perfect simplu, mai-mult-ca-perfect, persoanele 1-6
 }
subgrupa 3.1: {
identificator: *oac*;
alternanțe fonetice: *o -> oa* prezent, persoanele 1, 2, 6; imperfect, perfect simplu, mai-mult-ca-perfect, persoanele 1-6
 p -> c perfect simplu, mai-mult-ca-perfect, persoanele 1-6
 }
subgrupa 3.2: {
identificator: *oarc*;
alternanțe fonetice: *o -> oa* prezent, persoanele 1, 2, 6; imperfect, perfect simplu, mai-mult-ca-perfect, persoanele 1-6
 # -> *c* perfect simplu, mai-mult-ca-perfect, persoanele 1-6

}
 subgrupa 3.3: {
 identificator: *oat*;
 alternanțe fonetice: *o* -> *oa* prezent, persoanele 1, 2, 6; imperfect, perfect simplu, mai-mult-ca-perfect, persoanele 1-6
 ț -> *t* prezent, persoana 2
 }

subgrupa 4: {
 identificator: *ân*;
 alternanțe fonetice: *ă* -> *â* perfect simplu, mai-mult-ca-perfect, persoanele 1-6;
 # -> *n* prezent, persoana 2, perfect simplu, mai-mult-ca-perfect, persoanele 1-6
 }

subgrupa 5: {
 identificator: *un*;
 alternanță fonetică: *#* -> *n* prezent, persoana 2, perfect simplu, mai-mult-ca-perfect, persoanele 1-6
 }

subgrupa 6: {
 identificator: *up*;
 alternanță fonetică: nu există
 }

grupa 2: {
 subgrupa 1.1: {
 identificator: *rd*;
 alternanță fonetică: *z* -> *d* prezent, persoana 2
 }

subgrupa 1.2: {
 identificator: *rn*;
 alternanță fonetică: nu există
 }

subgrupa 2.1: {
 identificator: *at*;
 alternanțe fonetice: *ț* -> *t* prezent, persoana 2;
 ă -> *a* imperfect, perfect simplu, mai-mult-ca-perfect, persoanele 1-6
 }

subgrupa 2.2: {
 identificator: *ac*;
 alternanță fonetică: *ă* -> *a* imperfect, perfect simplu, mai-mult-ca-perfect, persoanele 1-6
 }

subgrupa 3: {
 identificator: *er, ec, ern, ep*;
 alternanță fonetică: nu există
 }

subgrupa 4.1: {
 identificator: *așt*;
 alternanțe fonetice: *ă* -> *a* imperfect, perfect simplu, mai-mult-ca-perfect, persoanele 1-6;
 }

sc -> *șt* prezent, persoana 1, perfect simplu, mai-mult-ca-perfect, persoanele
 1-6
 }
 subgrupa 4.2: {
 identificator: *eșt*;
 alternanță fonetică: *sc* -> *șt* prezent, persoana 1, perfect simplu, mai-mult-ca-perfect, persoanele
 1-6
 }
 subgrupa 5.1: {
 identificator: *oașt*;
 alternanțe fonetice: *o* -> *oa* prezent, persoanele 1, 2, 6; imperfect;
u -> *oa* perfect simplu, mai-mult-ca-perfect, persoanele 1-6
 }

 subgrupa 5.2: {
 identificator: *oas*;
 alternanțe fonetice: *o* -> *oa* prezent, persoanele 1, 2, 6; imperfect, perfect simplu, mai-mult-ca-
 perfect, persoanele 1-6
sc -> *șt* prezent, persoana 1, perfect simplu, mai-mult-ca-perfect, persoanele 1-
 6
 }
 subgrupa 6: {
 identificator: *in*;
 alternanțe fonetice: # -> *n* prezent, persoana 2
 }

În contextul acestei lucrări, verbele de conjugarea a IV-a sunt reunite în 9 grupe, dintre care 8 regulate și una neregulată, scrisă pentru verbul *a fi*. Clasele parțiale de terminații sunt în număr de 15:

a. patru dintre ele specifică timpurile perfect simplu și mai-mult-ca-perfect pentru grupele

- (1, 2, 3, 4, 7)- ConjIV.CPT 1;
- (5)- ConjIV.CPT 2;
- (6)- ConjIV.CPT 3;
- (8)- ConjIV.CPT 4;

b. trei cuprind terminații pentru timpul imperfect

- (1, 2, 7)- ConjIV.CPT 5;
- (3, 4, 8)- ConjIV.CPT 6;
- (5, 6)- ConjIV.CPT 7;

c. iar opt conțin terminații pentru timpul prezent

- (1)- ConjIV.CPT 8;
- (2)- ConjIV.CPT 9;
- (3)- ConjIV.CPT 10;
- (4)- ConjIV.CPT 11;
- (5)- ConjIV.CPT 12;
- (6)- ConjIV.CPT 13;
- (7)- ConjIV.CPT 14;
- (8)- ConjIV.CPT 15.

Identificatorii și liniile de alternanțe fonetice, dacă acestea există, sunt prezentate în continuare, în funcție de apartenența la una din cele 8 grupe verbale regulate:

grupa 1: {

identificator: *b, c, d, j, l, m, n, p, r, s, t, u, v, z, ver, găr*;

alternanță fonetică: nu există;

exemplu: a stabili, a citi, a primi, a privi, a vorbi, a găsi, a contribui, a munci, a gândi, a spori, a urzi, a vrăji, a adeveri

}

grupa 2: {

subgrupa1: {

identificator: *en*;

alternanțe fonetice: *i -> e* prezent, persoanele 1, 2, 3, 6;

n -> # prezent, persoana 2;

exemplu: a deveni

}

subgrupa2: {

identificator: *ș*;

alternanță fonetică: *s -> ș* prezent, persoanele 1, 3, 6;

exemplu: a ieși

}

subgrupa3: {

identificator: *ț*;

alternanță fonetică: *t -> ț* prezent, persoanele 1, 3, 6;

exemplu: a simți

}

subgrupa4: {

identificator: *orm*;

alternanțe fonetice: *oa -> o* prezent, persoana 3;

exemplu: a dormi

}

subgrupa 5: {

identificator: *g*;

alternanță fonetică: nu există;

exemplu: a fugi;

}

subgrupa6: {
 identificator: *ăr*;
 alternanță fonetică: *a* -> *ă* prezent, persoanele 1, 2, 3, 6;
 exemplu: a sări
 }
 subgrupa7: {
 identificator: *ur*;
 alternanță fonetică: *o* -> *u* prezent, persoanele 1, 2, 6; *oa* -> *u* prezent, persoana 3;
 exemplu: a sări
 }
 subgrupa8: {
 identificator: *os*;
 alternanțe fonetice: *ș* -> *s* prezent, persoana 2;
 oa -> *o* prezent, persoana 3;
 exemplu: a miroși
 }

grupa 3: {
 identificator: *ă*;
 alternanță fonetică: nu există;
 exemplu: a trăi
 }

grupa 4: {
 identificator: *itu, ibu, â*;
 alternanță fonetică: nu există;
 exemplu: a constitui, a contribui, a scârțâi
 }

grupa 5: {
 identificator: *or*;
 alternanță fonetică: *oa* -> *o* prezent, persoana 3;
 exemplu: a coborî
 }

grupa 6: {
 identificator: *ur*;
 alternanță fonetică: nu există;
 exemplu: a urî
 }

grupa 7: {
 identificator: *er*;
 alternanță fonetică: nu există;
 exemplu: a oferi
 }

grupa 8: {
 identificator: *șt*;
 alternanță fonetică: nu există;

exemplu: a ști
}

2.3.2 Obținerea formei verbale bază

Identificatorul de paradigmă este un subșir din forma verbală de infinitiv, iar baza de date morfologice este organizată pentru a include grupe de terminații, linii de alternanțe și restricții de unificare. Operațiile ce se aplică în interiorul regulilor lexicale sunt tot de tipul *scădere* și *adunare de șir*, respectiv de tip *alternanțe fonetice*. Validarea paradigmei parcurse se va realiza numai în condițiile găsirii în dicționar a formei lexicale obținute în finalul analizei.

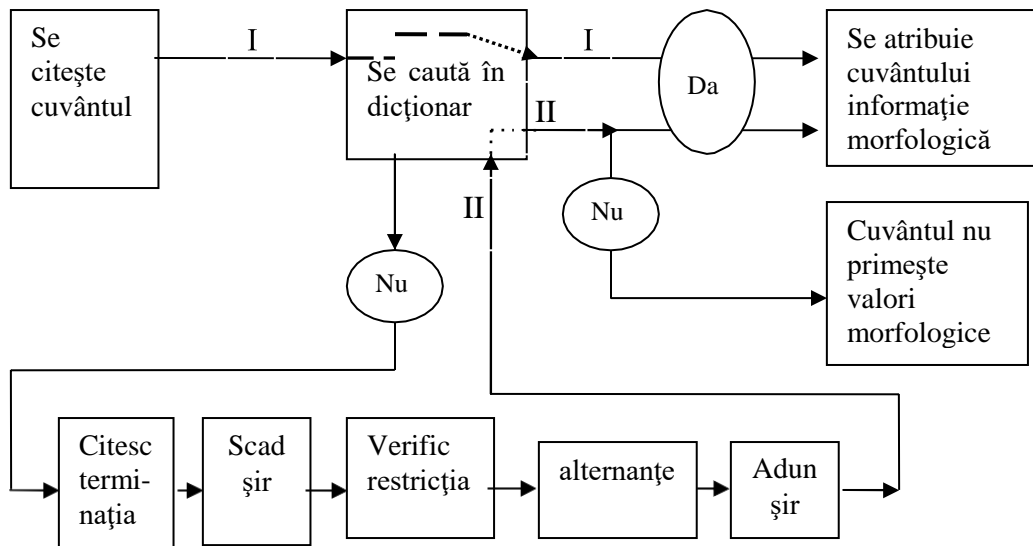


Figura 2.13: Schema logică de analiză a verbului

Pentru a înțelege mecanismul algoritmului de deflexionare bazat pe paradigmă verbală voi considera ca exemplu verbele cărora le corespunde ca restricție șirul *ăut*. Fie așadar forma verbală *căutasem*. Primul pas al algoritmului este căutarea în dicționar pentru a verifica dacă șirul de intrare corespunde unei intrări lexicale. Nu există această corespondență, iar algoritmul avansează spre pasul doi, citirea terminației. Pentru verb, am considerat terminații șirurile descrescătoare de maximum 8 caractere (numărul maxim de caractere pentru o terminație- d.e., *seseșăți*), în funcție de lungimea cuvântului (minimum 2 caractere). Se citește așadar, recursiv, șirul *utasem* de 6 litere, care nu se

regăsește în nicio grupă de terminații. Urmează șirul *tasem* cu aceeași finalitate ca mai sus, după care programul se oprește la terminația *asem*, identificată în clasa generală de terminații pentru verbele de conjugarea I.

În toate grupele de verbe de conjugarea I terminația *asem* corespunde timpului mai-mult-ca-perfect, persoanei I, singular. Se consideră, în ordine, prima grupă și se verifică dacă în tabelul de alternanțe fonetice există o astfel de linie care să respecte valorile morfologice obținute prin analiza terminației. Se găsește astfel alternanța: 1. ă<a imperfect; pf simplu; m.m.c.pf; 1, 2, 3, 4, 5, 6; grupa 11.

Cum însă identificarea grupei de terminații trebuie să fie în relație cu linia de alternanțe posibile, această linie este respinsă. Ea este asociată cu verbe din grupa 11, în consecință nu corespunde cu grupa terminației *asem*. Rezultă că nu există nicio linie de alternanțe, iar programul avansează spre pasul următor. După scăderea terminației *asem* din șirul de intrare, se verifică dacă rezultatul respectă restricția de identificare a grupei. Se obține șirul *căut* și se caută în tabelul de restricții subșirul cu care el se unifică. Este găsită restricția *ăut*, iar pasul este încheiat cu succes. Urmează altă etapă, cea a recompunerii intrării de dicționar, prin adăugarea morfemului de infinitiv. Deoarece terminația corespunde conjugării I, se va adăuga morfemul *a*, specific verbelor din această categorie. În acest moment, dacă este găsită intrarea lexicală *căuta*, paradigma este validată, în caz contrar algoritmul va trebui să reia toți pașii începând cu cealaltă grupă de verbe din lista identificată la început.

În aplicație, procesul de deflexionare verbală, are la bază trei elemente importante, definite ca obiecte separate, denumite Tgrup, Talt și Tres:

a. *Tgrup* reprezintă mulțimea grupelor de terminații în care sunt împărțite cuvintele. Am definit acest obiect ca având 4 variabile de tip string, ce reprezintă, în ordinea declarării lor, grupa de terminații, timpul verbului, persoana verbului și terminația corespunzătoare.

```
type Tgrup = class
public
    grupa, timp, pers, term: string;
    Constructor Create(g, t, p, m:string);
end;
```

O parte din paradigmele folosite în deflexionare sunt obiecte descendente din obiectul 'Tgrup', ele sunt păstrate pentru prelucrări ulterioare într-o listă numită 'lgrupe'. Pentru exemplificare luăm prima paradigmă, ce are următoarea sintaxă:

```
grupt111:=tgrup.Create('t1','prez','p1','');
```

unde grupt111 reprezintă codul de identificare al paradigmei, iar parametrii acesteia sunt: t1- grupa de terminații, prez- indică timpul prezent, p1- corespondent pentru persoana întâi, singular. Se observă că locul destinat valorii variabilei ce reprezintă terminația nu conține nimic, acest lucru fiind corect, deoarece această paradigmă este creată pentru forme ale verbelor în care nu există terminație care să aibă corespondența printre cele prestabilite (*merg – a merge*).

b. *Talt* a fost creat pentru a putea executa operații de tip alternanță, adică înlocuirea unor subșiruri din șirul de intrare cu alte șiruri. Datele membre acestui obiect corespund următoarelor elemente: grupa de terminații, timpul verbului, persoana verbului, setul 1 de alternanțe și setul 2 de alternanțe:

```
type talt = class
    public
        grupa, timp, alt11, alt12, alt21, alt22:string;
        pers:array[1..6] of string;
        constructor
create(g,t,p1,p2,p3,p4,p5,p6,a11,a12,a21,a22:string);
        procedure aplica(var s,a11,a12,a21,a22:string);
    end;
```

În acest caz, persoana verbului a fost declarată ca un vector de 6 variabile de tip string, deoarece există situații când un set de alternanțe corespund mai multor valori în același timp, sau chiar numărului total de valori pe care le poate avea această variabilă și astfel, putem declara toate valorile într-o singură linie de cod.

Fiind cazuri când unui cuvânt este nevoie să i se aplice două operații de tip alternanță pentru a ajunge la forma bază, am definit două seturi de alternanțe. Fiecare set conține două variabile, respectiv subșirul ce va fi căutat în șirul de intrare pentru

înlocuire, și șirul cu care acesta se va înlocui, dacă va fi găsit. Pentru exemplificare, folosim primul obiect descendent din obiectul 'Talt':

```
a1:= talt.create('t1','prez','p1','p2','p3',' ','p6','a','ț','ă','t');
lalt.Add(a1);
```

Paradigma se identifică prin codul 'a1' și este valabilă pentru grupa de alternanțe *tI*, timpul *prezent* și persoanele I, II și III singular, respectiv persoana a III-a plural. Seturile de alternanțe sunt compuse din (a, ț) și (ă, t). Toate paradigmele create ca obiecte descendente din *Talt* se păstrează în lista denumită *lalt*.

Obiectul *Talt* conține de asemenea și o metodă, numită *aplică*, prin intermediul căreia se realizează efectiv operația de tip alternanță.

Dacă subșirul corespunzător primei variabile din set nu se regăsește în forma identică în șirul de intrare, atunci operația de tip alternanță nu are loc. Raportat la exemplul de mai sus, principiul de funcționare devine: dacă variabilele *grupă*, *timpul* și *persoana*, conforme terminației găsite la un moment dat pentru cuvântul de analizat, sunt identice cu variabilele *grupă*, *timpul* și *persoana*, conforme paradigmei tip alternanța 'a1', atunci se vor căuta în șirul de intrare, de la dreapta la stânga, două subșiruri de un caracter, *a*, respectiv *ă*. În cazul identificării, se va trece la operația de înlocuire cu șirurile *ț* și/sau respectiv *t*.

c. *Tres* este un obiect cu rolul de a verifica existența unui anume subșir în șirul de intrare și, în urma unui rezultat pozitiv, va adăuga un alt șir (terminație) la sfârșitul acestuia, pentru obținerea lemei cuvântului de analizat. Subșirul în cauză reprezintă restricția paradigmei.

Obiectul *Tres* conține ca date membre două variabile de tip string denumite generic *sirdever* și *sirdead*, iar ca metodă, o procedură numită *adauga* ce realizează adăugarea terminației (*sirdead*) la sfârșitul cuvântului.

```
type tres = class
public
    sirdever, sirdead: string;
    constructor create (sv, sa:string);
    procedure adauga(var s, sa:string);
end;
```

Pentru exemplificarea modului de functionare vom considera prima instanță a obiectului, denumită 'r1':

```
r1:=tres.create('ăut','a');  
lres.Add(r1);
```

Prima variabilă a instanței *r1*, *ăut*, reprezintă șirul ce va fi căutat în interiorul șirului de intrare și, dacă acesta există, atunci șirul ce reprezintă cea de a doua variabilă a instanței, și anume 'a', va fi adăugat șirului de intrare, la sfârșitul acestuia. În caz contrar, se trece la verificarea următoarei instanțe. Toate paradigmele create ca restricții, adică instanțele obiectului *Tres*, vor fi introduse într-o listă special creată, denumită *lres*. Trebuie menționat faptul că restricțiile corespund anumitor grupe de terminații și au ca șiruri ce trebuie adăugate, forme bine prestabilite.

Capitolul 3

Flexiunea nominală și verbală în limba engleză

3.1 Flexiunea nominală în limba engleză

3.1.1 Problematizare

Limbajul și cunoașterea au fost explicate fie ca produse ale unei structuri omogene de memorie asociativă, fie ca un set de module de calcul determinate genetic, în care regulile manipulează reprezentări simbolice [74]. Distincția dintre lexicon și gramatică poate fi observată la nivelul subsistemului morfologic, între formele cu flexiune regulată și cele cu un comportament morfologic neregulat.

Flexiunea regulată este generată printr-o operație logică, determinată de reguli, incluse, la rândul lor, în constructe mai ample, sub forma paradigmelor nominale ori verbale. Flexiunea neregulată constă în stocarea și regăsirea articolelor lexicale în mod direct, fără apelul unei reguli de morfologie. Formele neregulate sunt astfel cuvinte stocate în lexicon, ce au și o trăsătură gramaticală fixată de tipul “past tense” sau “genitiv, singular, feminin”, ca descriere a intrării lexicale.

Într-un articol publicat în 2002, S. Pinker și M. Ullman argumentează teoria stocării formelor neregulate în lexicon, acesta fiind văzut ca o subdiviziune a memoriei declarative, responsabilă de evenimente, fapte și relații arbitrare, în timp ce formele regulate necesită existența unui sistem procedural.

Regularizarea lexicală nu este redusă la enunțuri declarative de tipul “pentru a crea timpul *past tense* în limba engleză, se adaugă verbului terminația *-ed*”, ci, așa cum se întâmplă, în morfologia paradigmatică, sunt relaționate o variabilă “V”, în cazul nostru intrarea lexicală a oricărui verb regulat și morfemul “-ed”, prin intermediul operației de unificare de șiruri.

Reprezentarea următoare este o ilustrare a teoriei lingvistice prin care se face diferență între cuvinte și reguli, între lexicon și gramatică: atunci când un cuvânt trebuie flexionat, lexiconul și gramatica sunt accesate în paralel. Dacă o formă flexionată pentru un substantiv sau un verb există în baza de date, așa cum este cazul formelor neregulate,

se trimite un semnal ce indică o legătură, oprindu-se căutarea. În caz contrar, procesorul gramatical generează o formă regulată prin aplicarea unei reguli de concatenare a două sau mai multe variabile morfologice:

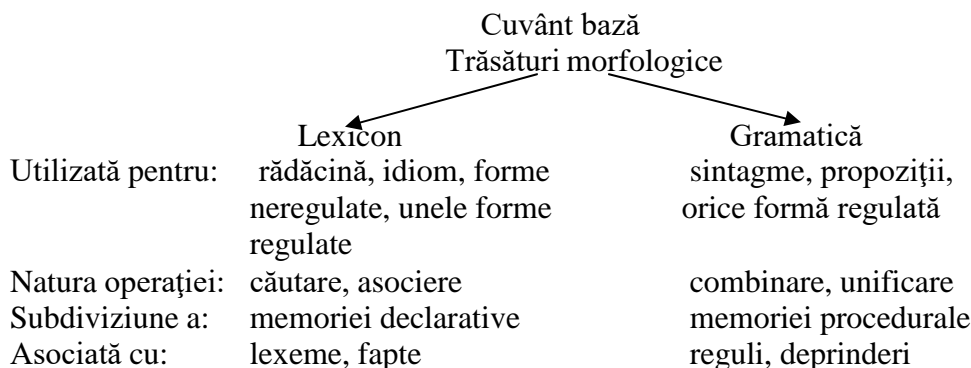


Figura 3.1: Ilustrare simplificată a teoriei Cuvinte-și-Reguli
(*Words and Rules Theory*, Pinker, 2002, p. 457)

Din mulțimea formelor neregulate, există unele care se comportă totuși în manieră regulată în anumite contexte, chiar dacă lexemul, în sine, nu poate fi considerat o intrare pentru sistemul flexionar:

- pentru limba română, de exemplu, substantivul „surori” are flexiune regulată pe traseul gramatical GD.fem.pl.art („surorilor”), NAc.fem.pl.art. („surorile”), NAc.fem.pl.neart. („surori”), GD.fem.sg.art. („surorii”);
- pentru limba engleză, de exemplu, verbele din contextele „ringed the city”, „steeled myself”, „bared his soul”, „costed out the grant” se comportă asemănător formelor regulate de *past tense*.

Unele din aceste regularizări sunt nesistematice- de exemplu, dubletele *dived / dove* și *dreamt / dreamed*, în care forma regulată e folosită sporadic, deoarece forma neregulată nu are o frecvență suficient de mare pentru a fi memorată. Multe regularizări sunt însă sistematice și constituie subiectul unor detalieri create de S. Pinker [75] pe baza unor experimente de laborator cu subiecți adulți și copii, astfel:

a. cuvinte fără rădăcină nominală sau verbală:

- onomatopoe: *dinged, pinged, zinged, peeped, beeped*;
- nume: *the Julia Childs, the Thomas Manns*;
- împrumuturi neasimilate: *talismans, mongooses*.

b. structura cuvintelor este exocentrică:

- verbe bazate pe nume, care sunt, la rândul lor, bazate pe verbe: *grandstanded, flied out, costed out the grant, encasted his leg*;

- substantive comune bazate pe nume, care sunt, la rândul lor, bazate pe substantive comune: *Mickey Mouses (simpletons), Renault Elfs, Top Shelves (frozen food), Seawolfs (aircraft)*;

- substantive ai căror referenți sunt diferiți de cei ai rădăcinii: *low-lifes, still lifes, sabre-tooths, Walkmans, tenderfoots*;

- substantive bazate pe sintagme: *Bag-A-Leafs, Shear-A-Sheeps*.

O altă regularizare sistematică apare pentru cuvintele derivate cu prefixe care își moștenesc trăsăturile de la morfemul cel mai din dreapta. De exemplu, rădăcina pentru „overeat” este „eat”; în consecință, „overeat” se referă la o formă anume de a mânca (deoarece moștenește trăsăturile semantice ale verbului „eat”) și are forma neregulată de *past tense*- „overate” (deoarece moștenește forma stocată de *past tense* a verbului „eat”). În limba română, verbul „a supralicita”, de exemplu, primește trăsăturile semantice (se referă la un fel de „a licita”) și cele de flexiune de la verbul „a licita”, care este, de altfel, și rădăcina acestuia.

Deși înțelesul formelor regulate diferă de cel al componentelor neregulate, regularizarea nu e condusă de rațiuni semantice. În cazul în care un cuvânt predispus a fi neregulat își păstrează înțelesul, dar păstrează, în același timp, și o rădăcină regulată în poziție centrală, atunci el rămâne tot neregulat:

- derivate cu prefix: *overate, overshoot, undid, preshrank, remade*;
- compunere: *bogeymen, superwomen, stepchildren, milkteeth*;
- metafore: *straw men, chessmen, snowmen, sawteeth, metrical feet, brainchildren, wolves in sheep's clothing*;
- idiomuri: *went out with (dated), went nuts (demented), went in for (chose), went off (exploded), went off (spoiled)*.

Pentru formele nominale și verbale regulate, redundanțele existente la nivelul datelor lingvistice pot fi utilizate pentru a acoperi generalizări pertinente, în condițiile în care datele sunt reprezentate prin paradigme gramaticale.

3.1.2 Schema logică de reprezentare a flexiunii nominale

Schema logică de reprezentare a flexiunii nominale operează cu modificări de șir asupra a două tipuri distincte de celule nominale: substantive regulate și substantive neregulate. Tipul regulilor de flexionare diferențiază între reguli aplicate numelor comune generale, neregulate și particulare (cu terminații de NAc. singular necesitând alternanțe fonetice și adăugări de șir diferite de morfemul de plural general). Formele de GD singular și plural sunt generate, pentru ambele categorii de nume comune, prin adăugări de șir, diferența fiind dată de generarea intermediară a cazului nominativ-acuzativ, plural.

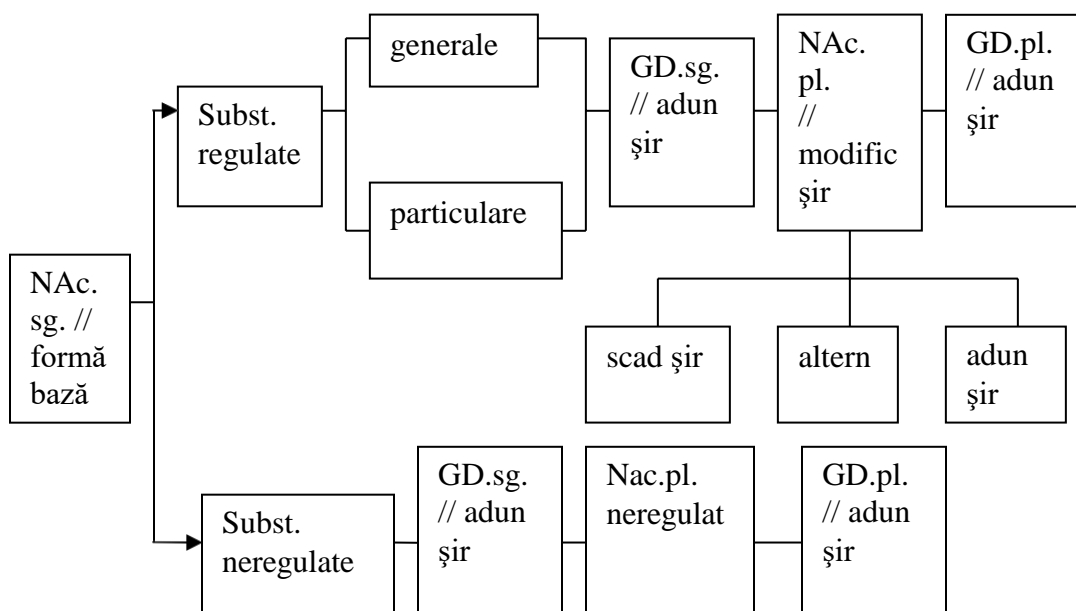


Figura 3.2: Schema logică de reprezentare a flexiunii nominale

3.1.3 Reguli de flexionare a numelui comun

a. Pluralul

Pentru cele mai multe substantive, forma de plural include morfemul **s** la sfârșitul lexemului: *dogs, trees, turtles*.

Substantivele terminate în **f, fe** suferă modificări de șir, prin alternanța fonetică **f** < **v** și adăugarea morfemului **es**. Tot o linie de alternanță fonetică suportă și substantivele cu terminație **y**, pentru care, la prima operație, **y** se transformă în **i**, iar la a doua operație se adaugă morfemul de plural **es**. Numele comune cu una din terminațiile **ch, o, sh, x, s, z, ss** primesc marca de plural **es**: *beaches, foxes, wishes*.

Excepțiile de la aceste reguli de flexionare sunt introduse în dicționar cu forme de singular și plural (d.e., *cliff – cliffs, roof – roofs, proof – proofs; studio – studios, piano – pianos*).

Substantivele cu o formă de plural trimitând la înțelesuri diferite au câte două intrări de dicționar, una pentru înțelesul de singular comun cu cel de plural, și a doua pentru înțelesul diferit față de forma de singular:

Custom – obicei

Customs – obiceiuri; vamă

Manner – mod

Manners – moduri; comportament

Spectacle – spectacol

Spectacles – spectacole; ochelari

Formele nominale neregulate, introduse ca atare în dicționar și apelate prin memoria lexicală, generează pluralul asemenea unor părți de vorbire lipsite de flexiune: *child – children, man – men, woman – women, goose – geese, mouse – mice, tooth – teeth, ox – oxen, die – dice, foot – feet, louse – lice*.

b. Genitivul sintetic

Numele comun singular formează genitivul sintetic prin adăugarea unui apostrof urmat de morfemul **s**: *the car's engine*. În schimb, substantivele terminate în **s**, **ss**, **z** sau **x** omit morfemul de genitiv.

Numele comune de plural adaugă ‘**s** dacă nu se termină în **s**, iar în caz contrar, primesc doar apostroful: *children' books, singers' voices*.

c. Genitivul analitic

Genitivul analitic este creat după regula *obiect posedat + of + posesor*, iar utilizarea lui este generică pentru lucruri, insecte, păsări mici, când substantivul provine din adjectiv sau e precedat de un determinat.

d. Dativul

Dativul este construit prepozițional, prin luarea în considerație a relației de dependență sintactică între prepoziția *to* și grupul nominal cu formă substantivală de singular sau de plural.

3.2 Conjugarea verbului în limba engleză

3.2.1 Introducere

Verbele pot fi clasificate din punctul de vedere al structurii morfologice sau din punctul de vedere al sensului lexical și al funcției [40].

Din punctul de vedere al structurii morfologice, clasificarea verbelor se realizează în modul următor:

- a. verbe simple: *come*;
- b. verbe compuse: *spotlight*;
- c. verbe cu particulă adverbială: *put on, put off, put away*;
- d. verbe cu prepoziție obligatorie: *look at, listen to, wait for*;
- e. locuțiuni verbale: *take care of, make use of*.

Din punctul de vedere al sensului lexical, verbele din limba engleză se împart în:

- a. verbe cu sens lexical plin, numite, de obicei, noționale;

b. verbe cu sens lexical redus, folosite mai mult cu funcții gramaticale.

În plan sintactic, verbele cu sens lexical plin sunt predicative, iar verbele cu sens lexical redus sunt nepredicative și au nevoie de o complinire (verb noțional, nume predicativ) pentru a deveni predicate. În schimb, ele marchează anumite categorii gramaticale.

În clasa verbelor cu sens lexical redus intră verbele copulative, auxiliare și modale:

a. verbele copulative: *be* și, conform gramaticii tradiționale, *appear*, *seem*, *look*, *turn*, *become*, *get*, *remain*, *stay*, etc.: *He seems happy*.

b. verbele auxiliare: *be*, *have*, *do*, *shall*, *will*, care îndeplinesc funcția de marcă a categoriei gramaticale de diateză, mod, timp, aspect, persoană și număr, a formei interogative și negative la verbele pe care le însoțesc.

c. Verbele semiauxiliare sau modale: *can*, *must*, *may*, care arată atitudinea vorbitorului față de enunț.

Verbele lexicale și verbele auxiliare diferă atât prin morfologia flexionară, cât și prin sintaxă. Mulțimea formelor flexionare ale unui lexem variabil constituie o paradigmă, iar toate verbele lexicale au, în principiu, o paradigmă cu șase forme, ilustrată astfel:

Forme	Mood and tense	Person and number	Take	Want	Hit
Primary	Preterite		Took	Wanted	Hit
	Present tense	3 rd sg	Takes	Wants	Hits
		plain	Take	Want	Hit
Secondary	Plain form		Take	Want	Hit
	Gerund participle		Taking	Wanting	Hitting
	Past participle		Taken	Wanted	Hit

Figura 3.3: Paradigma verbelor lexicale

Pentru propozițiile enunțative propriu-zise, cu predicatie afirmativă, paradigmele verbale vor include:

- terminații dependente de persoană, număr și cel mai din dreapta fonem al verbului lexical;
- grupa auxiliarului utilizat pentru timpurile compuse și aspectul continuu;

- o linie de alternanțe fonetice aplicate persoanei a 3-a singular, prezent, dependente de vocala internă sau ultimul fonem al cuvântului de intrare.

Pentru exemplificare, propunem grupa verbală de terminații T1, unde terminație se consideră unul din șirurile de caractere: *s, z, sh, x, ch, o*. Formele auxiliarelor Be1, Have1, Be2, Be3, Have2, Be4 sunt specificate în josul tabelului de reprezentare a formării verbelor predicative, la modul indicativ.

Timpul	Pers.		Altern.	Aux.1	Aux.2
	3	1,2,4,5,6			
Present Tense Simple	es	-	-		
Present Tense Continuous	ing	ing	-	Be1	
Past Tense Simple	ed	ed	-		
Past Tense Continuous	ing	ing	-	Be2	
Present Perfect Simple	ed	ed	-	Have1	
Present Perfect Continuous	ing	ing	-	Have1	Be3
Past Perfect Simple	ed	ed	-	Have2	
Past Perfect Continuous	ing	ing	-	Have2	Be3
Future Tense Simple	-	-	-	1, 4- Shall 1,2,3,4,5,6 - Will	
Future Tense Continuous	ing	ing	-	Be4	
Future Perfect Simple	ed	ed	-	1- Shall 1,2,3,4,5,6 - Will	Have
Future-in-the-Past Simple	-	-	-	1- Should 1,2,3,4,5,6 - Would	

Tabela 3.1: Conjugarea verbelor din grupa de terminații T1 la modul indicativ

Be1: {1 – am; 2, 4, 5, 6 – are; 3 – is};
 Be2: {1, 3 – was; 2, 4, 5, 6 – were}
 Have2: {have – had}
 Have1: {1, 2, 4, 5, 6 – have; 3 – has}
 Be3: {be - been}
 Be4: {1 – shall be; 1, 2, 3, 4, 5, 6 – will be}

3.2.2 Schema logică de conjugare a verbului

Schema logică de reprezentare a conjugării verbului în limba engleză cuprinde ca unități celulare distincte: forma bază și cea obținută prin conjugare, auxiliarele verbale, clasele de verbe diferențiate prin terminații generale ori specifice și operațiile de modificare de șiruri (*scad șir*, *adun șir*, *alternez șiruri*).

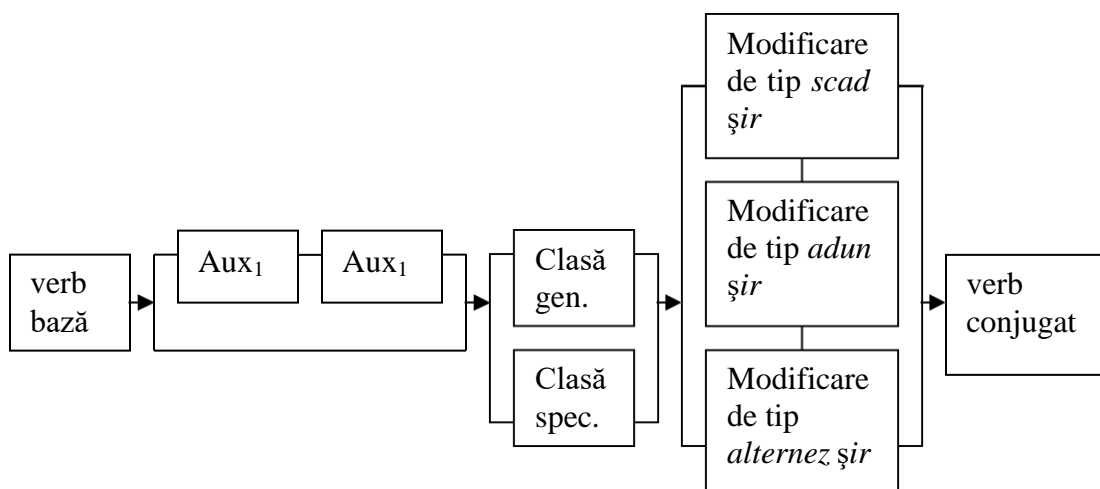


Figura 3.4: Schema logică de reprezentare a conjugării verbului

Pașii algoritmului de conjugare pornesc de la citirea formei verbale bază, echivalenta intrării de dicționar. Dacă verbul trebuie să ajungă la un timp compus cu auxiliar sau la aspectul continuu, se adaugă înaintea acestuia forma auxiliară corespunzătoare descrierii din tabelul de reguli. Operația specifică unei clase verbale generale este exclusiv de adăugare de șir, iar pentru clasele verbale specifice, diferențiate prin terminații, sunt activate toate operațiile posibile de modificare: *scad*, *adun șir*, respectiv *alternez șiruri*. După realizarea ultimei operații incluse în reprezentarea fiecărei

reguli de conjugare, se obține o formă verbală cu informații gramaticale de mod, timp, persoană, număr, indicate atât prin auxiliar(e), cât și prin terminațiile verbale.

3.2.3 Obținerea formei verbale conjugate

Pentru cele mai multe verbe, indicativul prezent de persoana a 3-a singular se creează prin adăugarea morfemului *s* la forma de infinitiv. Totuși, pentru verbele terminate în: {*s, z, sh, ch*} e necesară adăugarea morfemului *es*, iar pentru verbele cu structură fonetică finală {*consoană + y*} se creează, mai întâi, alternanța fonetică *y -> i*, apoi șirul astfel rezultat este modificat prin adăugarea subșirului *es*. Cele trei auxiliare au comportament flexionar diferit, rezultând următoarele subclase verbale de prezent:

Be1: {1 – am; 2, 4, 5, 6 – are; 3 – is};

Have1: {1, 2, 4, 5, 6 – have; 3 – has};

Do1: {1, 2, 4, 5, 6 – do; 3 – does}.

Verbele regulate, conjugate la Past Tense Simple, pot fi categorizate, în funcție de comportamentul flexionar, în patru clase:

- a. cu fonemul *e* ca terminație de infinitiv și primind morfemul *d*;
- b. cu structura fonetică dată de [*consoană + y*] și primind morfemul *ed*;
- c. cu structura fonetică dată de [*vocală + consoană ≠ {w, y}*] și dublând consoana finală, înainte de a primi morfemul *ed*;
- d. generală, care primește la forma de infinitiv morfemul *ed*.

La rândul lor, și verbele neregulate aparțin de trei categorii principale. Categorizarea lor însă nu construiește tipare flexionare, deoarece formele neregulate au un comportament impredictabil, care determină imposibilitatea de a le genera automat, pe baza unor reguli. Deși se pot recunoaște unele șabloane pentru verbele neregulate, ele rămân rezistente formalizării prin reguli, purtând mărcile lingvistice și neuropsihologice ale memoriei lexicale²⁰:

²⁰ Pinker, Steven, [2000]. *The Irregular Verbs*. Landfall, March, 2000

- a. verbe care nu comportă nicio schimbare fonetică la trecerea de la forma de infinitiv la cea de *past tense simple*: {de exemplu, beat – beat; cost – cost; fit – fit; let – let; read – read; set – set; shut – shut; spread – spread};
- b. verbe care schimbă vocala internă ori cea finală: {de exemplu, begin – began; come – came; dig – dug; draw – drew; drink – drank; drive – drove; fall – fell; forget – forgot; give – gave; grow – grew; hold – held; know – knew; lead – led; meet – met; ring – rang; rise – rose; run – ran; shine – shone; sing – sang; slide – slid; stride swim – swam};
- c. verbe care schimbă consoana finală: {de exemplu, build – built; burn – burnt; deal – dealt; dream – dreamt; fell – felt; hear – heard; learn – learnt; mean – meant; smell – smelt; spell – spelt; spend – spent};
- d. verbe care se schimbă complet în timpul conjugării: {de exemplu, bear – bore; break – broke; bring – brought; buy – bought; catch – caught; eat – ate; fight – fought; find – found; fly – flew; forbid – forbade; go – went; keep – kept; lay – laid; leave – left; lie – lay; lose – lost; make – made; pay – paid; say – said; see – saw; seek – sought; sell – sold; show – showed; sleep – slept; speak – spoke; stand – stood; take – took; teach – taught; tell – told; think – thought; understand – understood};
- e. Be2: {1, 3 – was; 2,4,5,6 – were};
- f. Have2: {have – had}
- g. Do2: {do – did}.

Pentru formele de Present Perfect Simple, clasificarea verbelor neregulate se realizează păstrând, în linii mari, aceleași criterii precum cele specificate formelor de Past Tense:

- a. verbe care nu comportă nicio schimbare fonetică la trecerea de la forma de infinitiv la cea de *present perfect simple*: {burst – burst; come – come; cost – cost; cut – cut; hit – hit; let – let; put – put; read – read; run – run; set – set; shut – shut; spread – spread};
- b. verbe care schimbă vocala internă ori cea finală: {begin – begun; dig – dug; draw – drew; drink – drunk; feed – fed; hang – hung; hold – held; lead – led; meet –

- met; ring – rung; shine – shone; shoot – shot; shrink – shrunk; sing – sung; sink – sunk; slide – slid; stick – stuck; swim – swum};
- c. verbe care schimbă consoana finală: {bend – bent; build – built; burn – burnt; deal – dealt; dream – dreamt; fell – felt; hear – heard; learn – learnt; mean – meant; smell – smelt; spell – spelt; spend – spent};
- d. verbe care se schimbă complet în timpul conjugării: {blow – blown; bring – brought; buy – bought; catch – caught; draw – drawn; fight – fought; find – found; fly – flown; go – gone; grow – grown; keep – kept; leave – left; lose – lost; make – made; pay – paid; say – said; seek – sought; sell – sold; show – shown; sleep – slept; sow – sown; stand – stood; swear – sworn; sweep – swept; teach – taught; tell – told; think – thought; understand – understood};
- e. verbe care adaugă morfemul *en* la forma de infinitiv: {beat – beaten; broke – broken; choose – chosen; drive – driven; eat – eaten; fall – fallen; forbid – forbidden; forget – forgotten; freeze – frozen; give – given; hide – hidden; ride – ridden; rise – risen; see – seen; snake – snaken; speak – spoken; steal – stolen};
- f. Be3: {be – been}
- g. Have3: {have – had}
- h. Do3: {do – done}

Auxiliarul *to be* folosit în crearea viitorului comportă două forme, în funcție de persoană și număr:

Be4: {1 – shall be; 1, 2, 3, 4, 5, 6 – will be}.

3.2.4 Grupele de verbe din limba engleză, diateza activă

Investigând transformările pe care le parcurge verbul regulat în limba engleză, diateza activă, am obținut cinci grupe verbale, clasificate după terminația specifică de infinitiv și după modificările pe care rădăcina verbală le înregistrează pe parcursul conjugării. Cele cinci grupe verbale sunt prezentate în continuare, pentru fiecare în parte existând și un exemplu care să ilustreze comportamentul întregii grupe:

T1. terminația {s, z, sh, x, ch, o} // ex.: pass -- passes

T2. terminația {consoană + y} // ex.: study – studies

T3. terminația {vocală + consoană}, unde:

- a. înaintea vocalei nu mai există o altă vocală;
- b. consoana este diferită de {w,x,y};
- c. verbul este monosilabic (maximum 4 caractere) // ex.: to nod – nodding

T4. terminația {e} // ex.: to delete deleting

T5. terminație generală.

Generarea formei verbale regulate afirmative, la diateza activă, începe prin căutarea verbului în lista de verbe neregulate. Dacă există, se echivalează informația gramaticală din limba română în limba engleză și se extrage forma corespunzătoare. Altfel, se identifică grupa verbului, în funcție de terminațiile de la infinitiv (grupele T1-T5), se echivalează informația gramaticală din română în engleză și se compune forma corespunzătoare, conform descrierilor din tabelul de mai jos:

Timpul	Clasă verb	Pers.		Altern.	Aux. ₁	Aux. ₂
		3	1,2,4,5,6			
Present Tense Simple	T3, T4, T5	s	-	-		
	T1	es	-	-		
	T2	es	-	y -> i		
Present Tense Continuous	T1, T2, T5	ing	ing	-	Be1	
	T3	[+ cons.] ing ²¹	[+ cons.] ing	-	Be1	
	T4	[-e]ing ²²	[-e]ing	-	Be1	
Past Tense Simple	T1, T5	ed	ed	-		
	T4	+d	+d	-		

²¹ Dublează consoana, înainte de a primi terminația specifică aspectului continuu, *ing*.

²² Elimină fonemul *e*, înainte de a primi terminația *ing*.

	T2	ed	ed	y -> i		
	T3	[+ cons.] ed	[+ cons.] ed	-		
Past Tense Continuous	T1, T2, T5	ing	ing	-	Be2	
	T3	[+ cons.] ing	[+ cons.] ing	-	Be2	
	T4	[-e]ing	[-e]ing	-	Be2	
Present Perfect Simple	T1, T5	ed	ed	-	Have1	
Present Perfect Simple	T4	+d	+d	-	Have1	
	T2	ed	ed	y -> i	Have1	
	T3	[+ cons.] ed	[+ cons.] ed	-	Have1	
Present Perfect Continuous	T1, T2, T5	ing	ing	-	Have1	Be3
	T3	[+ cons.] ing ²³	[+ cons.] ing	-	Have1	Be3
	T4	[-e]ing ²⁴	[-e]ing	-	Have1	Be3
Past Perfect Simple	T1, T5	ed	ed	-	Have2	
	T4	+d	+d	-	Have2	
	T2	ed	ed	y -> i	Have2	
	T3	[+ cons.] ed	[+ cons.] ed	-	Have2	
Past Perfect Continuous	T1, T2, T5	ing	ing	-	Have2	Be3
	T3	[+ cons.] ing ²⁵	[+ cons.] ing	-	Have2	Be3
	T4	[-e]ing	[-e]ing	-	Have2	Be3
Future Tense Simple	T1, T2, T3, T4, T5	-	-	-	1- Shall 1,2,3,4,5,6- Will	

²³ Dublează consoana, înainte de a primi terminația specifică aspectului continuu, *ing*.

²⁴ Elimină fonemul *e*, înainte de a primi terminația *ing*.

²⁵ Dublează consoana, înainte de a primi terminația specifică aspectului continuu, *ing*.

Future Tense	T1, T2, T5	ing	ing	-	Be4
Continuous	T3	[+ cons.] ing	[+ cons.] ing	-	Be4
	T4	[-e]ing	[-e]ing	-	Be4

Tabela 3.2: Modul indicativ, forma afirmativă: verb_{afirmativ}

Timpurile simple și compuse ale modului indicativ sunt create, pentru forma negativă, utilizând negația **not**. Valorile gramaticale de persoană și număr sunt incluse în marca auxiliarului, care devine purtătorul principal de informație verbală. Am păstrat cele cinci grupe de verbe regulate descrise la forma afirmativă, aceleași operații de modificare a șirului de intrare prin adăugare, ștergere de șir sau alternanță de șiruri, iar variațiile auxiliarelor sunt specificate în maniera următoare:

a. prezent:

DO_{1Neg}: {1, 2, 4, 5, 6 – do not / don't; 3 – does not / doesn't}Be_{1Neg}: {1, 4 – am not; 2, 5, 6 – are not / aren't; 3 – is not / isn't}Have_{1Neg}: {1, 2, 4, 5, 6 – have not / haven't; 3 – has not – hasn't}

b. past tense:

DO_{2Neg}: {1, 2, 3, 4, 5, 6 – did not / didn't}Be_{2Neg}: {1, 3 – was not / wasn't; 2, 4, 5, 6 – were not / weren't}Have_{2Neg}: {1, 2, 3, 4, 5, 6 – had not / hadn't}

c. participiu trecut:

Be: {been}

Have₃: {had}Do₃: {done}

d. viitor:

Be_{4Neg}: {1, 4 – shall not be / shan't be; 1, 2, 3,4,5,6 – will not be / won't be}

Timpul	Clasă verb	Persoană	Alt. fon.	Aux.1	Aux.2
Present Tense Simple	T1, T2, T3,			DO _{1Neg}	
	T4, T5				
Present	T1, T2, T5	ing		Be _{1Neg}	

Tense Continuous	T3	[+ cons.] ing		Be _{1Neg}	
	T4	[-e]ing		Be _{1Neg}	
Past Tense Simple	T1, T2, T3, T4, T5			Do _{2Neg}	
Past Tense Continuous	T1, T2, T5	ing		Be _{2Neg}	
	T3	[+ cons.] ing		Be _{2Neg}	
	T4	[-e]ing		Be _{2Neg}	
Present Perfect Simple	T1, T5	ed		Have _{1Neg}	
	T4	+d		Have _{1Neg}	
	T2	ed	y -> i	Have _{1Neg}	
	T3	[+ cons.] ed		Have _{1Neg}	
Present Perfect Continuous	T1, T2, T5	ing		Have _{1Neg}	Be ₃
	T3	[+ cons.] ing		Have _{1Neg}	Be ₃
	T4	[-e]ing		Have _{1Neg}	Be ₃
Past Perfect Simple	T1, T5	ed		Have _{2Neg}	
	T4	+d		Have _{2Neg}	
	T2	ed	y -> i	Have _{2Neg}	
	T3	[+ cons.] ed		Have _{2Neg}	
Past Perfect Continuous	T1, T2, T5	ing		Have _{2Neg}	Be ₃
	T3	[+ cons.] ing		Have _{2Neg}	Be ₃
	T4	[-e] ing		Have _{2Neg}	Be ₃
Future Tense Simple	T1, T2, T3, T4, T5			1, 4- Shall not / shan't 1,2,3,4,5,6- Will not / won't	
Future Tense Continuous	T1, T2, T5	ing		Be _{4Neg}	
	T3	[+ cons.] ing		Be _{4Neg}	
	T4	[-e]ing		Be _{4Neg}	

Tabela 3.3: Modul indicativ, forma negativă: verb_{negativ}

3.3. Descrierea implementării flexiunii nominale și verbale

Flexionarea, în limba engleză, se realizează atât pentru verbe, cât și pentru substantive, dar în module separate, care deși au anumite principii de funcționare asemănătoare, folosesc tipuri diferite, dar specifice de proprietăți ale cuvintelor. Ca asemănări, menționez că ambele module au ca element principal câte un obiect numit *tflexvb*, respectiv *tflexsb*, ce conține tipurile specifice de proprietăți ale cuvintelor necesare procesului de flexionare (de exemplu, pentru verbe avem *timp* și *persoană* , iar pentru substantive avem *caz* și *număr*) și elemente auxiliare (alternanțe, terminații, particule). Ambele module se folosesc de terminațiile formei bază a cuvântului în limba țintă pentru a stabili grupa din care acesta face parte, grupă care, ulterior, va indica diverse operații ce se vor executa asupra cuvântului de analizat.

3.3.1 Conjugarea verbului

Conjugarea verbului se bazează pe un număr de cinci grupe de terminații pentru care s-au creat 90 de paradigme de flexionare. Paradigmele sunt de tipul obiectului *tflexvb*, amintit anterior, obiect ce este definit astfel :

```
type tflexvb = class
    public
        egrupa, etimp, ecodtimp, estgterminatie, eadgterminatie, eauxiliar1, eauxiliar2,
        ealternanta1,
        ealternanta2:string;
        epers:array[1..6] of string;
        constructor create (eg, et, ect, ep1,ep2,ep3,ep4,ep5,ep6,estgt,eadgt,ea1,ea2,
        ealt1,ealt2:string);
end;
```

unde :

- egrupa(eg) reprezintă grupa de terminații;
- etimp(et) reprezintă timpul verbului (valoarea acestei variabile rezultă din codul morfologic) ;

- ecodtimp(ect) reprezintă proprietatea de aspect pe care o au verbele în limba engleză și poate avea două valori : ‘c’ pentru *continuu* și ‘s’ pentru *simplu*;
- epers(ep1..ep6) reprezintă persoana verbului; spre deosebire de celelalte variabile, aceasta este definită ca un vector de șiruri, astfel am putut restrânge numărul paradigmelor care conțineau operații identice pentru persoanele diferite ale aceluiași cuvânt, într-o singură paradigmă, în care se specifică toate persoanele corespunzătoare;
- estgterminatie(estgt) reprezintă terminația formei bază ce va fi ștearsă;
- eadgterminatie(eadgt) reprezintă terminația ce se va adăuga formei bază;
- eauxiliar1(ea1) reprezintă primul verb auxiliar cu ajutorul căruia se creează timpurile compuse;
- eauxiliar2(ea2) reprezintă al doilea verb auxiliar cu ajutorul căruia se creează timpurile compuse pentru aspectul continuu al verbului lexical; în cazul în care avem o structură cu două auxiliare, cel corespunzător variabilei eauxiliar1, în practică, îl va precede pe cel corespunzător variabilei eauxiliar2;
- ealternanta1(ealt1) reprezintă subșirul ce va fi șters din șirul de intrare conform operației de tip alternanță;
- ealternanta2(ealt2) reprezintă șirul ce va fi adăugat în locul subșirului șters anterior pentru a încheia operația de tip alternanță;

La începutul procedurii, după ce se inițializează lista ‘lflexvb’, listă în care se vor adăuga, pe rând, toate paradigmele necesare flexionării, se vor declara aceste paradigme. Ca exemplu, luăm paradigma denumită ‘f15’, deoarece are o structură mai încărcată în valori, acest lucru favorizând clarificarea construcției paradigmei:

```
f15:=tflexvb.create('e2','pesi','s','p1','p2','p4','p5','p6',"",'ed','have',"','y','i');
lflexvb.add(f15);
```

Acest exemplu poate fi interpretat după cum urmează:

- cuvintele care corespund acestei paradigme sunt verbe ce fac parte din grupa numărul 2 de terminații,
- verbul este la timpul perfect simplu,

- corespondența aspectului verbal în engleză este ‘*simple*’,
- următoarele șase locații sunt rezervate declarării persoanelor verbului; aceste locații se completează numai cu valorile valide pentru paradigma curentă; dacă numărul valorilor este mai mic de 6, atunci restul de locații rezervate variabilei *epers* rămân goale; persoanele verbului, în cazul de față, pot fi I și a II-a singular, precum și I, a II-a și a III-a plural;
- deoarece locația ‘estgt’ nu conține nicio valoare, înseamnă că nu avem terminație de șters;
- locația variabilei *eadgterminatie* conține valoarea ‘ed’, deci la sfârșitul cuvântului va fi adăugat șirul respectiv;
- locația ‘ea1’ conține valoarea ‘have’, acest lucru însemnând că verbul va fi precedat de această particulă;
- locația *ea2* nu conține nicio valoare, deci avem o structură ce se creează cu un singur auxiliar, și anume cel corespunzător variabilei anterioare *eauxiliar1*;
- ultimele două locații ne arată că această paradigmă presupune și existența operației de tip alternanță în care subșirul reprezentat de caracterul ‘y’ va fi înlocuit cu șirul, tot de un singur caracter, și anume ‘i’.

După declararea paradigmelor, procesul continuă prin inițializarea variabilelor corespunzătoare proprietăților morfologice (exemplu: *ts* – reprezintă timpul verbului, *prs* – reprezintă persoana verbului) și auxiliare (exemplu: *codts* – reprezintă aspectul verbului în limba engleză) ale cuvântului de flexionat. Valorile acestor variabile sunt extrase din codul morfologic ce însoțește cuvântul respectiv. În continuare se verifică condițiile de stabilire a grupei de terminații în care se încadrează forma bază a cuvântului în limba țintă, după care începe procedura de căutare a paradigmei corespunzătoare.

Validarea unei paradigme se realizează numai în momentul în care patru din proprietățile verbului, grupa, timpul, aspectul verbului și persoana, coincid cu cele ale paradigmei. După identificarea paradigmei valide, în funcție de grupa în care a fost încadrată forma bază a verbului în limba țintă și de existența valorilor pentru variabilele de tip șterge_terminație, adaugă_terminație, auxiliar1, auxiliar2 și alternanță, în structura paradigmei se efectuează modificările necesare flexionării cuvântului respectiv. Când

setul de modificări stabilit a fost efectuat, se finalizează și procesul de flexionare pentru verb, iar valoarea obținută se adaugă șirului ce reprezintă traducerea finală.

3.3.2 Flexiunea nominală

Flexiunea nominală implică aceiași pași principali ca și cei folosiți în conjugarea verbului, deosebirea dintre cele două module fiind dată de tipul variabilelor. Obiectul utilizat pentru a defini structura paradigmelor create pentru flexionarea substantivului se numește ‘tflexsb’. Acesta este definit astfel :

```
type tflexsb = class
  public
    grupa, animat, caz, numar, articol, alternanta1, alternanta2, terminatie,
    auxiliar1,
    auxiliar2 :string;
    constructor create (grs, anims, cazs, nrs, arts, as1, as2, tes, aus1,
    aus2:string);
  end;
```

unde :

- grupa(grs) reprezintă grupa de terminații pentru lemele substantivelor în limba țintă; aceste grupe sunt în număr de 6, iar criteriile de categorizare se bazează pe condiții referitoare la terminațiile formelor bază nominale;
- animat(anims) reprezintă trăsătura semantică *animat* atribuită substantivelor; această variabilă primește valoarea ‘a’ dacă substantivul face referire la o entitate animată;
- caz(cazs) reprezintă cazul substantivului;
- număr(nrs) indică numărul substantivului;
- articol(arts) precizează dacă substantivul este articulat sau nu;
- alternanta1(aş1) reprezintă subșirul care se va șterge din șirul de intrare în cadrul operației de tip alternanță;
- alternanta2(aş2) reprezintă șirul care se va adăuga în șirul de intrare în locul celui șters pentru a încheia operația de tip alternanță;

- terminație(tes) este șirul care se va adăuga la sfârșitul șirului de intrare cu rol de terminație, indicând pluralul sau genitivul sintetic;
- auxiliar1(aus1) este șirul ce se va adăuga formei substantivale, înaintea ei și separat de aceasta, indicând cazul dativ sau genitiv perifrastic;
- auxiliar2(aus2) reprezintă articolul definit proclitic; în cazul unei structuri alcătuite din două auxiliare, particula de caz precede particula de articol.

Voi exemplifica modul de utilizare al acestui obiect cu ajutorul paradigmei denumite 'fs1e1' :

```
fs1e1:=tflexsb.create('g1','a','d','p','a','f','v','es','to','the');
lflexsb.add(fs1e1);
```

Interpretarea asociată paradigmei exemplificate este următoarea :

- substantivele care folosesc această paradigmă pentru flexionare se află în grupa 1 de terminații;
- trăsătura semantică are valoarea *a*, adică entitatea referită de substantiv este animată; în cazul în care trăsătura semantică nu are completată nicio valoare, se consideră ca entitatea este inanimată;
- substantivul ce corespunde acestei paradigme se află în cazul dativ;
- trăsătura de număr are valoarea *plural*;
- substantivul este articulat hotărât;
- următoarele două valori indică faptul că paradigma necesită operația de tip alternanță, astfel încât subșirul *f* va fi înlocuit cu șirul *v*;
- valoarea *es* a variabilei terminației va fi adăugată la sfârșitul formei substantivale;
- variabila auxiliar1 are valoarea *to*, specificând informația cazuală de dativ; ea va precede substantivul, separat de acesta;
- ultima valoare menționată în paradigmă, *the*, este articolul hotărât proclitic; aceasta va avea o poziție intermediară între informația cazuală de dativ și substantiv.

În continuare, procesul de flexionare nominală se desfășoară analog procesului de flexionare verbală descris anterior. Se inițializează variabilele ce reprezintă trăsăturile morfologice și semantice, cu valori extrase din codul morfologic al cuvântului; se

stabilește grupa de terminații din care face parte lema; se caută paradigma care să îndeplinească condițiile de validare, după care se execută modificările specificate în vederea generării formei flexionate. Aceasta va fi adăugată șirului ce reprezintă traducerea finală.

Capitolul 4

Modulul de preprocesare a textelor TTL

4.1 Prezentare generală

Modulul de preprocesare a textelor TTL²⁶ este disponibil ca serviciu web implementat atât în varianta SOAP [98], cât și REST [99], ce produce, următoarele tipuri de adnotări²⁷:

A. **Recunoașterea entităților denumite**²⁸: TTL utilizează expresii regulate pentru identificarea tipurilor de entități (nume proprii de persoane, de țări, de instituții, date, numere reale, întregi, etc.), iar fiecare expresie regulată are asociată o etichetă ce specifică tipul entității.

Un expert codifică într-o gramatică nerecursivă care suportă operatorii de repetiție Kleene {*, +, ?} câte o regulă pentru fiecare tip de entitate. Regulile sunt apoi traduse automat prin expandarea lor în expresii regulate Perl și se impune asupra acestora un filtru care să verifice ordinea lor de aplicare. Expresiile regulate ce recunosc șiruri mai lungi de caractere primesc prioritate mai mare față de restul expresiilor.

B. **Segmentarea la nivel de frază**²⁹: TTL folosește o serie de șabloane pentru identificarea sfârșitului de frază și o listă de abrevieri frecvente pentru o limbă dată.

Odată cu recunoașterea entităților denumite se elimină situațiile în care segmentarea s-ar fi putut realiza în interiorul unei entități care ar fi conținut un simbol de sfârșit de frază. Șabloanele de sfârșit de frază rezolvă situațiile în care apare punctuația pereche: dacă după un simbol de sfârșit de frază există o paranteză închisă sau ghilimele închise sau apostrof închis, atunci acestea trebuie păstrate în fraza curentă.

²⁶ În engleză, “Tokenizing, Tagging and Lemmatizing free running texts”.

²⁷ Actualmente, serviciul web prelucrează limbile română, engleză și franceză, dar implementarea TTL este independentă de limbă și poate fi extinsă în măsura în care resursele lingvistice folosite de componenta de învățare sunt disponibile.

²⁸ În engleză, “Named Entity Recognition”.

²⁹ În engleză, “Sentence Splitting”.

Lista de abrevieri frecvente într-o limbă dată (TTL utilizează, pentru limba română, 731 de abrevieri, iar pentru limba engleză 186) este folosită astfel: punctul final după un cuvânt din lista de abrevieri marchează finalul abrevierii, iar dacă după punctul final urmează un cuvânt care începe cu inițială majusculă, atunci punctul este, de asemenea, și final de frază.

C. **Segmentarea la nivel de cuvânt**³⁰: folosește liste de expresii pe care le recunoaște în text și liste de prefixe și sufixe care, dacă fac parte dintr-un cuvânt, sunt despărțite de acesta.

Sunt considerate unități lexicale de sine stătătoare acele expresii idiomatice al căror înțeles este nedecompozabil și pentru care spațiul nu mai reprezintă un marcaj de sfârșit de cuvânt. Cum însă există situații în care o secvență de cuvinte poate fi expresie într-un context, iar în altul nu- TTL utilizează o listă de secvențe de cuvinte care, indiferent de context, constituie expresii și sunt recunoscute ca atare într-o frază (Figura 1):

mai_cu_seamă	COMPOUND
peste_poate	COMPOUND

Figura 4.1: Compuși românești ca unități lexicale

În situațiile în care dintr-o secvență de caractere ce nu conține spațiu trebuie extrase două sau mai multe cuvinte, șirul vid este considerat, din perspectiva modulului TTL, un separator de cuvinte. În acest scop a fost creată o listă de prefixe și sufixe care trebuie separate dacă sunt identificate într-un șir de caractere care nu conține spațiu, această listă precizând astfel pozițiile în care șirul vid este separator de cuvânt (Figura 2):

într-	LEFTSPLIT	prepoziție
-o	RIGHTSPLIT	pronume

Figura 4.2: Prefixe (LEFTSPLIT) și sufixe (RIGHTSPLIT) care trebuie separate în română

³⁰ În engleză, “Tokenizing”.

D. **Adnotarea cu etichete morfosintactice**³¹: o etichetă morfosintactică este o codificare a unei părți de vorbire împreună cu valorile asociate variabilelor ce o descriu. TTL implementează tehnica etichetării stratificate [99], o metodă originală ce permite modelarea statistică a dezambiguizării de etichete morfosintactice în termenii unor inventare conținând un număr foarte mare de etichete.

TTL folosește două inventare de etichete morfosintactice aflate în corespondență: primul inventar de etichete respectă specificațiile MULTEXT-EAST (caz în care eticheta se numește MSD- din engleză, “Morpho-Syntactic Descriptor”), iar al doilea inventar este derivat din primul eliminându-se din fiecare etichetă morfosintactică variabilele morfosintactice care nu sunt dependente de context (eticheta derivată se numește CTAG- din engleză, “Corpus TAG”).

TTL implementează adnotatorul cu etichete morfosintactice TnT⁷ îmbunătățit cu câteva euristici noi. Acesta este un adnotator probabilistic bazat pe Modele Markov Ascunse (MMA, “Hidden Markov Models”, în engleză). Folosește un corpus adnotat pentru a-și estima probabilitățile de tranziție, iar stările automatului sunt trigrame de etichete morfosintactice CTAG.

În cazul cuvintelor necunoscute, implementarea TTL diferă de cea originală TnT prin:

- analiza de sufix, adică atribuirea unei etichete morfosintactice unui cuvânt pe baza analizei ultimelor sale caractere;
- dacă lexemul începe cu majusculă și nu se află la început de frază, adnotatorul are opțiunea de a-l eticheta ca fiind substantive propriu;
- dacă recunoașterea entităților denumite a fost realizată într-un pas anterior de prelucrare a textului, entitățile au primit deja etichetă morfosintactică.

E. **Lematizarea**³²: utilizează un model de leme extras automat dintr-un lexicon în care fiecare ocurență are asociate o leamnă și eticheta morfosintactică proprie.

³¹ În engleză, “Part-of-Speech Tagging”.

³² În engleză, “Lemmatization”.

Lematizarea este o operație de normalizare a formei ocurente a unui cuvânt, operație care transformă orice formă ocurentă a unui cuvânt într-o formă standard, adoptată, de regulă, de dicționare³³. TTL se bazează pe un model de lematizare extras automat dintr-un lexicon care conține pentru fiecare formă ocurentă a unui cuvânt, lema și MSD-ul acestuia (care poate fi transformat în CTAG-ul corespunzător).

Algoritmul de lematizare primește la intrare o listă de cuvinte, fiecare dintre ele având asociat CTAG-ul său. Dacă S este o listă de perechi formă ocurentă, CTAG $\langle w_i, t_i \rangle$, algoritmul de lematizare parcurge pașii următori³⁴:

1. dacă forma ocurentă a cuvântului w_i este întâlnită în lexicon, extrage lema din intrarea respectivă folosind cheia $\langle w_i, t_i \rangle$; dacă lema nu este unică, alege cea mai frecventă lema dintr-o listă de frecvențe pentru leme.
2. aplică toate regulile de lematizare specifice etichetei t_i și obține o listă de leme candidate. Ordonează descrescător această listă după probabilitățile furnizate de modelul Markov pentru eticheta t_i și extrage lema din capul listei. Deoarece modelul Markov tinde să aleagă lema cu lungimea cea mai mică, în modulul TTL a fost introdusă în procesul de selecție și frecvența regulii care a produs lema respectivă.

F. Analiza sintactică de suprafață a constituenților gramaticali³⁵

Algoritmul de analiză identifică grupuri frazale atomice (NP, VP, AP, PP). Pentru fiecare atom lexical se atașează în reprezentarea XCES atributul *chunk* specificând grupul / grupurile frazale din care face parte.

4.2 Utilizarea modulului de preprocesare a textelor TTL

În cadrul aplicației AGV1.0, modulul pentru preprocesare a textelor TTL este folosit pentru performanțele obținute în lematizare. Înainte de a detalia cum anume este

³³ R. Ion, **Metode de dezambiguizare semantică automată. Aplicații pentru limbile engleză și română**, 2007

³⁴ ibidem

³⁵ În limba engleză, „Chunking”

folosit în aplicația AGV 1.0, facem o incursiune în modul de funcționare al acestuia folosind exemplul de mai jos.

Considerăm textul literar extras din *Sărmanul Dionis*, de M. Eminescu, pentru care modulul TTL creează următoarele adnotări:

1. text de intrare:

„Trecut și viitor e în sufletul meu, ca pădurea într-un sâmbure de ghindă, și infinitul asemenea, ca reflectarea cerului înstelat într-un strop de rouă.”

2. aspect gramatical urmărit: dezambiguizarea morfologică

3. rezultat TTL:

```
<w lemma="trecut" ana="Ncms-n" chunk="Np#1">Trecut</w>
<w lemma="și" ana="Crssp" chunk="Np#1">și </w>
<w lemma="viitor" ana="Ncms-n" chunk="Np#1">viitor</w>
<w lemma="fi" ana="Vmip3s" chunk="Vp#1">e</w>
<w lemma="în" ana="Spsa" chunk="Pp#1">în</w>
<w lemma="suflet" ana="Ncmsry" chunk="Pp#1,Np#2">sufletul</w>
<w lemma="meu" ana="Ds1ms-s" chunk="Pp#1,Np#2">meu</w><c>,</c>
<w lemma="ca" ana="Rc">ca</w>
<w lemma="pădure" ana="Ncfsry" chunk="Np#3">pădurea</w>
<w lemma="întru" ana="Spsay" chunk="Pp#2">într-</w>
<w lemma="un" ana="Timsr" chunk="Pp#2,Np#4">un</w>
<w lemma="sâmbure" ana="Ncms-n" chunk="Pp#2,Np#4">sâmbure</w>
<w lemma="de" ana="Spsa" chunk="Pp#3">de</w>
<w lemma="ghindă" ana="Ncfsrn" chunk="Pp#3,Np#5">ghindă</w>
<c>,</c>
<w lemma="și" ana="Crssp">și</w>
<w lemma="infinit" ana="Ncmsry" chunk="Np#6">infinitul</w>
<w lemma="asemenea" ana="Afp" chunk="Np#6,Ap#1">asemenea</w>
<c>,</c>
<w lemma="ca" ana="Rc">ca</w>
<w lemma="reflectare" ana="Ncfsry" chunk="Np#7">reflectarea</w>
<w lemma="cer" ana="Ncmsoy" chunk="Np#7">cerului</w>
<w lemma="instela" ana="Vmp--sm" chunk="Vp#2">înstelat</w>
```

```

<w lemma="întru" ana="Spsay" chunk="Pp#4">într-</w>
<w lemma="un" ana="Timsr" chunk="Pp#4,Np#8">un</w>
<w lemma="strop" ana="Ncms-n" chunk="Pp#4,Np#8">strop</w>
<w lemma="de" ana="Spsa" chunk="Pp#5">de</w>
<w lemma="rouă" ana="Ncfsrn" chunk="Pp#5,Np#9">rouă</w>
<c>.</c>

```

4. comentariu: programul TTL dezambiguiază ocurențele cu multiple valori morfosintactice și reține o singură valoare pentru fiecare lexem, în parte (de exemplu, lexemul *de* este analizat ca prepoziție simplă, marcă a cazului acuzativ, valoarea morfosintactică fiind selectată înaintea aceluia de conjuncție subordonatoare sau interjecție³⁶; lexemul *și* poate fi adverb sau conjuncție, dar în adnotarea făcută de TTL cuvântul este dezambiguit cu valoarea de conjuncție coordonatoare simplă).

Dezambiguizarea morfologică reprezintă procesul care stabilește o valoare morfologică unică pentru o parte de vorbire ambiguă.

În cadrul studiului prezentat în această lucrare, problema dezambiguizării morfologice este rezolvată într-un mod interactiv: dacă în timpul procesării se constată că termenului curent t_i i s-au atribuit mai multe valori morfosintactice, programul afișează un mesaj cu două opțiuni:

1. fie urmează să se afișeze valorile morfosintactice respective, utilizatorul trebuind să o indice pe cea corectă;
2. fie se deschide automat un program creat de Institutul de Cercetare pentru Inteligența Artificială (I.C.I.A.), care folosește serviciile web TTL. În fereastra acestui program se introduce textul de analizat sau se indică un fișier de unde acesta să fie extras. Apoi, programul pregătește datele în formatul TTL, se conectează la serverul ICIA³⁷, accesează modulul TTL descris mai sus și preia rezultatele pe care le scrie într-un fișier specificat de utilizator. În continuare, aplicația AGV1.0 preia din fișierul respectiv valoarea morfosintactică unică atribuită termenului curent t_i . În aceste condiții procesul de dezambiguizare morfologică

³⁶ În variantă populară, cuvântul *de* poate fi și pronume: DE pron. rel. invar. (Pop.) Care, ce. Omul de-l văzu. (<http://dexonline.ro/definitie/de>)

³⁷ <http://www.racai.ro>

pentru acest termen ia sfârșit. În cazul în care mai există și alte leme ambigue în fraza de analizat, aplicația caută automat valoarea morfosintactică unică în fișierul deja creat, ce conține analiza textului, efectuată de modulul de preprocesare a textelor TTL.

După stabilirea valorilor morfosintactice pentru toți termenii frazei de analizat, programul avansează la modulul de flexiune nominală și verbală pentru limba engleză.

Capitolul 5

Sintaxa grupului verbal pentru limbile română și engleză

5.1 Tipuri de gramatici generative

În studiul **Gramaticile generative**, E. Vasiliu [101] consideră două metodologii lingvistice diferite de abordare a studiului sincron al limbii: una analitică și alta sintetică. Analiza are la bază ideea că limba este o structură care poate fi descompusă în unitățile componente, iar rezultatul analizei este un inventar de unități (foneme, morfeme, cuvinte, etc.) și o clasificare a acestor unități. Sinteza pleacă de la identificarea posibilităților și restricțiilor combinatorii ale elementelor rezultate din analiză și își propune să obțină reprezentarea structurii limbii.

Descrierea structurii limbii a devenit posibilă în momentul în care s-a adoptat și în lingvistică metoda modelării: “Modelul folosit pentru descrierea structurii limbii este schema de funcționare a unei mașini cibernetice construită în așa fel încât, cu ajutorul unui ansamblu finit de simboluri ale cuvintelor împreună cu informația gramaticală legată de fiecare cuvânt, ansamblu care poartă numele de vocabular sau alfabet, să poată produce toate propozițiile corecte ale unei limbi și numai aceste propoziții.”³⁸ Deoarece modelul construit este o mașină care produce propoziții, modelarea descrisă a structurii gramaticale poartă numele de gramatică generativă³⁹.

Cel mai simplu model generativ se referă la gramaticile cu un număr finit de stări. Acestea modelează descrierea lingvistică în termenii unui dispozitiv de tipul lanțurilor Markov, care trecând de la o stare la alta, produce un simbol (un cuvânt, un morfem, un fonem, etc.) la fiecare tranziție.⁴⁰ O propoziție este construită prin alegeri succesive: un vorbitor alege din mulțimea cuvintelor unul pe care îl enunță; pentru construirea în continuare a propoziției alege un al doilea, apoi un al treilea, ș.a.m.d. Alegerile sunt condiționate de conținutul semantic al comunicării și de anumite restricții în succesiunea cuvintelor.

³⁸ Emanuel Vasiliu, **Gramaticile generative**, Limba română, nr. 3, 1963, pp. 219-231

³⁹ http://ebooks.unibuc.ro/filologie/dominte/10-2.htm#_ftn1

⁴⁰ E. Vasiliu, Sanda Golopenția Eretescu, **Sintaxa transformățională a limbii române**, 1969, pp. 15-37

Gramaticile cu un număr finit de stări nu gestionează însă situații gramaticale de tipul propozițiilor intercalate sau al anticipării obiectului direct prin clitic pronominal. În propoziția *Ea îl vede pe Sabin*, alegerea de către vorbitor a termenului *îl* nu este impusă de alegerea precedentă (termenul *Ea*), după cum nici alegerea prepoziției *pe* nu este condiționată de alegerea verbului *vede*. Ambii termeni (*îl*, *pe*) trebuie raportați la substantivul propriu *Sabin*, care, ca obiect direct, cere antepunerea prepoziției *pe* și dublarea pronominală.

Un al doilea model generativ pornește de la teoria constituenților imediați. Prin segmentări succesive și atribuirea calității de subordonat termenului substituibil cu zero, gramatica bazată pe teoria constituenților imediați prezintă generarea unui enunț ca pe un proces continuu de descompunere a unor unități (constituenți) de rang superior în unități de rang inferior.

Printre încercările de sintaxă structurală, metoda analizei în constituenți imediați se deosebește de analiza sintactică tradițională prin faptul că operează cu diviziuni succesive⁴¹, iar rezultatul analizei sunt niște segmente numite constituenți imediați. Criteriul de bază al segmentării enunțului în constituenți imediați îl constituie substituirea: se încearcă înlocuirea unei părți din enunț cu alt material lingvistic, astfel încât rezultatul să fie acceptat de normele limbii respective. În urma tăierii enunțului în constituenți imediați, trebuie rezolvată problema care din diviziunile posibile rezultate este cea mai bună. În acest moment se alege pentru fiecare diviziune creată un focar și un context. Acea parte din enunț care rămâne neschimbată în cursul operațiilor de substituire se numește context, iar partea de enunț asupra căreia intervin încercările de substituire se numește focar. Se studiază bogăția clasei de focare⁴² și se cercetează variația contextului care admite fiecare clasă de focare. În enunțul *Papagalul mănâncă mere*, clasa de focare care admite reprezentanți ai claselor de secvențe de tipul: *papagalul*, *el*, *pasărea*, etc.,

⁴¹ “De obicei, enunțul se împarte mai întâi în două, fiecare din cele două segmente se împarte iarăși în două, și așa mai departe, până se ajunge la segmente care nu mai pot fi divizate, deoarece s-ar obține complexe sonore fără înțeles.”- Solomon Marcus, Edmond Nicolau, Sorin Stati, *Introducere în lingvistica matematică*, vol. I, 1966, pp. 26-41

⁴² “Se alege acea segmentare din care rezultă o clasă de distribuție mai bogată în: a) numărul de contexte în care poate apărea și b) numărul claselor de secvențe pe care le înglobează. În enunțul *El cumpără o pălărie*. dacă se constată, de pildă, că clasa de focare a lui *el* e mai bogată decât clasa de focare a lui *el cumpără o*, segmentarea propoziției se face după *el*, acesta fiind prin urmare primul constituent imediat al enunțului” – Sorin Stati, op. cit.

poate apărea în contexte ca: *mănâncă verdeață, cântă vesel, zboară*. Finalizarea acestor operații conduce la următoarele concluzii: diviziunea care creează clasele de forme cele mai bogate este cea mai bună dintre toate diviziunile posibile; cele două părți rezultate din segmentarea enunțului sunt constituenții imediați ai acestuia.

Pentru a afla ce relație există între părțile componente ale unui enunț se aplică testul substituirii unuia dintre constituenții enunțului cu zero. În enunțul *Cartea galbenă este interesantă*, înlocuirea cu zero a părții *galbenă* conduce la un enunț admis de limbă, în timp ce substituirea constituentului *cartea* cu zero are ca finalitate un enunț nereperat de limbă: **galbenă este interesantă*. În concluzie, dacă nu obținem un enunț reperat, partea omisă este considerată supraordonată sau regentă (*cartea*, în enunțul de mai sus), iar când substituirea este acceptată, partea omisă este subordonată⁴³ (constituentul *galbenă*, în enunț, este subordonat față de *cartea*).

Modelul bazat pe constituenții imediați prezintă avantajul de a fi ușor de transformat în algoritm și de folosit în diverse aplicații practice bazate pe sinteza automată a limbilor. Față de modelul anterior, al gramaticilor cu număr finit de stări, orientat spre raportul de succesiune a termenilor în lanț, acest model exprimă în mod explicit raporturile de diverse naturi existente între unitățile lingvistice (cuvinte, morfeme, etc.). În același timp, față de următorul model, cel al gramaticilor transformaționale, are ca dezavantaj faptul că nu poate explica parafrazele (echivalențe, sinonimii) sintactice. Pentru gramatica de constituenți imediați, construcția *citirea cărții* reprezintă un grup nominal format din [nume + nume în genitiv]. Pentru gramatica transformațională, construcția *citirea cărții* este o transformare a propoziției *citește cartea*.

Modelul gramaticii transformaționale se bazează pe premisa că structura unei limbi poate fi descrisă sub forma unui număr finit de transformări ale unor structuri de bază, mai simple, numite propoziții nucleare.⁴⁴ Gramatica transformațională își propune astfel două obiective fundamentale:

⁴³ În loc de subordonat-regent, se preferă în lingvistica americană perechile de termeni satelit-nucleu sau modificator-centru.

⁴⁴E. Vasiliu, op. cit.

- a. generarea tuturor enunțurilor corecte dintr-o limbă și exclusiv a acestora.⁴⁵
- b. descrierea structurală a enunțurilor generate.⁴⁶

Regulile de transformare se pot defini ca reguli care convertesc o secvență cu o anumită structură a constituenților într-o secvență cu o structură diferită a constituenților, creându-se următoarea tipologie:

- reguli prin care un constituent al unei propoziții este înlocuit cu o propoziție (de exemplu, regulile care generează fraze prin subordonare);
- reguli prin care o propoziție este *nominalizată*, altfel spus este transformată în construcție cu verb la mod nepersonal și introdusă drept constituent al unei alte propoziții (de exemplu, regulile care generează construcțiile gerunziale, participiale, infinitivale, etc.);
- reguli prin care structura unei propoziții este modificată, fără a se face apel la structura unei alte propoziții (de exemplu, regulile care generează propoziții pasive).

Dintre cele trei modele generative descrise, gramatica de constituenți imediați și gramatica transformațională au o putere explicativă mai mare decât gramaticile cu un număr finit de stări, deoarece acestea din urmă permit doar reprezentarea structurii limbii la un singur nivel.

În continuarea lucrării **Naturaletă limbaului artificial – studiu despre traducerea automată a GV**, am ales modelul gramaticii de constituenți imediați pentru a reprezenta relațiile la nivelul grupului verbal. În etapa de cercetare am identificat tipul constituenților grupului verbal în limbile română și engleză, precum și relațiile dintre aceștia. Exemplele asupra cărora m-am oprit în analiza sintactică ilustrează stilul beletristic pe care l-am considerat mai apropiat de flexibilitatea și naturaletă limbii prin diferitele dispuneri sintactice ale constituenților GV. Fiecare exemplu este reprezentat sintactic cu ajutorul programului Linguistic Tree Constructor, disponibil pe pagina de internet a organizației Summer Institute of Linguistics⁴⁷. De asemenea, fiecare exemplu

⁴⁵ Realizarea acestui obiect conferă gramaticii o capacitate generativă slabă (*weak generative capacity*).- cf. N. Chomsky (1965) **Aspects of the theory of syntax**, The MIT Press, Cambridge

⁴⁶ Realizarea acestui obiectiv asigură gramaticii o capacitate generativă puternică (*strong generative capacity*). – cf. N. Chomsky, 1965

⁴⁷ www.sil.org

este însoțit de tiparul verbal asociat, cel pe care l-am folosit ulterior în etapa de implementare pentru crearea regulilor și pentru evaluarea lor

5.2 Linguistic Tree Constructor

În gramatica tradițională românească, inventarul unităților sintactice oscilează de la două (propoziția și fraza), la trei (partea de propoziție, propoziția și fraza) și la patru [partea de propoziție, grupul de cuvinte (sintagma), propoziția și fraza]. Sintaxa românească structurală înlocuiește propoziția și fraza cu „enunțul”, pe care-l consideră unitatea sintactică de bază [88], [44], [48]. Extinzând inventarul unităților sintactice, Vasile Șerban adaugă paragraful și textul [89].

Partea de propoziție reprezintă cea mai mică unitate sintactică, exprimată printr-o parte de vorbire funcțională care participă la realizarea unei comunicări. O propoziție este compusă din părți de propoziție, dintre care unele actualizează singure o funcție sintactică (*Vine toamna.*), iar altele actualizează o funcție sintactică împreună cu un auxiliar care poate fi: morfologic: *El va veni; A terminat de scris*; sintactic: *Vine la Universitate. Acest medicament se folosește în caz de intoxicație*; sau semantic: *A venit numai Ion; Nici azi nu vine.*

Reprezentarea cea mai frecventă a unei propoziții este cea a arborelui sintactic, numit- în general- indicator sintagmatic.

Formal, fiind dată o gramatică independentă de context, un arbore sintactic este un arbore cu următoarele proprietăți:

1. rădăcina este etichetată prin simbolul inițial;
2. fiecare frunză este etichetată printr-o unitate lexicală sau prin ε ;
3. fiecare nod interior este etichetat printr-un non-terminal;
4. dacă A este non-terminalul etichetând nodul interior și etichetele fiilor acestui nod sunt, de la stânga la dreapta, X_1, X_2, \dots, X_n , atunci $A \rightarrow X_1, X_2, \dots, X_n$ este o regulă de producție.⁴⁸ Un caz particular îl constituie regula $A \rightarrow \varepsilon$, care semnifică faptul că un nod etichetat A are un singur fiu, etichetat ε .

⁴⁸ X_1, X_2, \dots, X_n reprezintă fie un non-terminal, fie un terminal.

Fiecare nod are o etichetă care corespunde unei sintagme sau unei categorii. Sintagmele reprezintă o proiecție, în timp ce categoriile sunt noduri terminale. Fiecărui nod terminal îi este asociat un cuvânt (sau un articol lexical).

Linguistic Tree Constructor (LTC)⁴⁹ este o aplicație pe care am utilizat-o în această lucrare pentru crearea arborilor sintactici. Versiunea LTC 3.0.3. este o versiune Unicode, funcționabilă pentru platformele: Windows, Mac Os X, Linux/Unix.

LTC asistă expertul uman în construirea arborilor sintactici dintr-un text, aceștia putând fi de tipul: generic, X-Bar sau al Gramaticii de Rol și Referință⁵⁰ (RRG).

Pentru arborii generici, ierarhia cuprinde, de jos în sus: cuvântul, sintagma și nodul sintactic. Tipurile “Clause”, “Sentence” și “Paragraph” nu pot fi editate de utilizator, în schimb celelalte noduri sintactice pot fi definite în funcție de scopul analizei. Există două fișiere de biblioteci localizate diferit. Mai întâi, programul caută în directorul de configurare pe care utilizatorul îl poate crea și controla singur. Dacă fișierul este găsit, căutarea se oprește. În caz contrar, programul caută într-un director controlat de program.

La deschiderea programului LTC, apare o fereastră împărțită în șase suprafețe, detaliate în continuare (Figura 5.1):

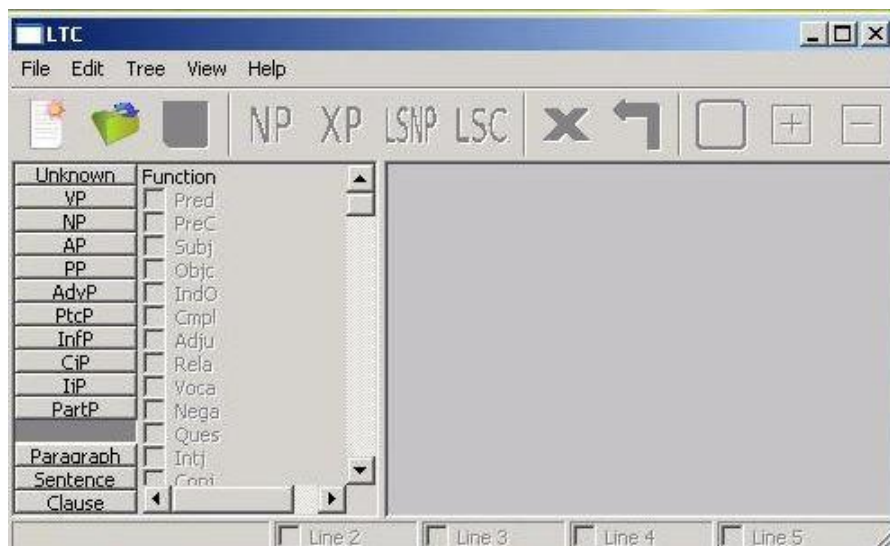


Figura 5.1: Fereastra de deschidere a programului LTC

1. Partea principală a ferestrei, unde sunt creați și afișați arborii sintactici;

⁴⁹ disponibil la adresa: <http://ltc.sourceforge.net/>

⁵⁰ În engleză, “Role and Reference Grammar”.

2. Partea cea mai din stânga- aceasta conține un număr de butoane pentru crearea sintagmelor și celorlalte noduri;
3. Partea din mijloc- pentru selectarea etichetelor asociate unui nod;
4. Bara de instrumente- de unde majoritatea comenzilor programului sunt activate;
5. Bara-status, de unde poate fi selectat modul de dispunere interlinear (Line 2, Line 3, Line 4, Line 5);
6. Bara-menu.

Partea principală a ferestrei este locul în care utilizatorul interacționează cel mai mult cu programul, aceste interacțiuni fiind de tipul: expandare / colapsare nod, selectare / deselectare nod, selectare a unui nod simultan cu deselectarea tuturor celorlalte, selectarea unei ierarhii a nodurilor.

Ca exemplificare a acestei interacțiuni de tip utilizator - program, voi prezenta modul în care am intervenit în programul LTC pentru a obține reprezentarea sintactică a două haiku-uri⁵¹ în limbile română și engleză (Ex. 1):

Exemplul 1:

<i>În zborul lor,</i>	<i>În their flight,</i>
<i>vulturii redesenează</i>	<i>eagles redraw</i>
<i>spiralele scoicilor.</i>	<i>the shells volutes.</i>

Pentru crearea arborilor sintactici generici, am editat fișierul de etichete asignat nodurilor, adăugând etichete de tip parte de vorbire⁵², iar pentru etichetele *funcție* am introdus obiectul direct (DirO). Am construit arborii sintactici, folosind comenzile din bara de instrumente din stânga și din mijlocul ferestrei principale, arbori care pot fi parcurși pe verticală sau pe orizontală (Figura 5.2):

⁵¹ Florin Vasiliu, *Tolba cu licurici*, Editura Haiku, București, 1993, p. 40

⁵² Părțile de vorbire introduse sunt aceleași din codificarea standard utilizată pe tot parcursul lucrării și respectă specificațiile Multext-East. Pentru partea de vorbire codificată în specificațiile Multext-East drept *Adposition*, am introdus două valori morfosintactice- *Sp*, pentru prepoziție, *St*, pentru postpoziție. Pentru verb, am introdus etichetele V, Va și Vo, ultimele două specificând valorile de verb auxiliar și verb modal.

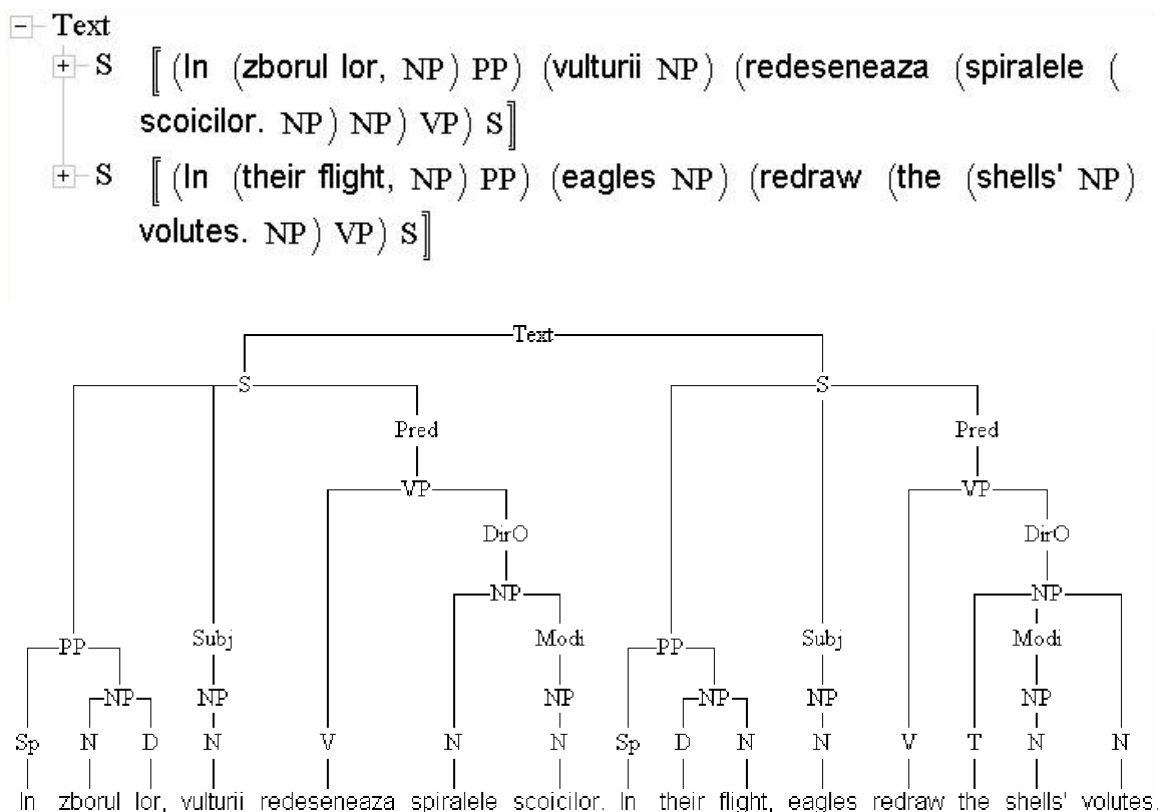


Figura 5.2: Deschiderea arborilor sintactici pe verticală (prima figură) și pe orizontală (a doua figură) pentru exemplul 1

5.3 Etapa de cercetare. Verbele bivalente

Pentru a desemna relațiile care se stabilesc între procesul exprimat de verb și participanții implicați de verb se folosește termenul *valență*.

Verbele bivalente sunt caracterizate prin combinarea cu două vecinătăți actanțiale obligatorii, cu valori semantice și forme diferite ale actanților [1]. Structurile bivalente includ valența de stânga, primul actant, subiectul, și valența de dreapta, al doilea actant, complementul direct.

Schema celulară a grupului verbal tranzitiv pornește de la detalierea structurii subclaselor verbelor bivalente, pentru care o valență obligatorie este satisfăcută de subiect. Deoarece relația de inerență dintre subiect și predicat este tratată la un alt nivel

decât cel de generalizare a structurii verbale, voi include în schema celulară exclusiv verbul și argumentele interne. Paradigma de flexiune verbală cuprinde formele verbului lexical pentru modurile predicative, la diateza activă.

Schema generală de reprezentare a relației dintre verbul – predicat și argumentul intern include legătura existentă între obiectul direct și formele pronominale clitice prin care obiectul este anticipat sau reluat:

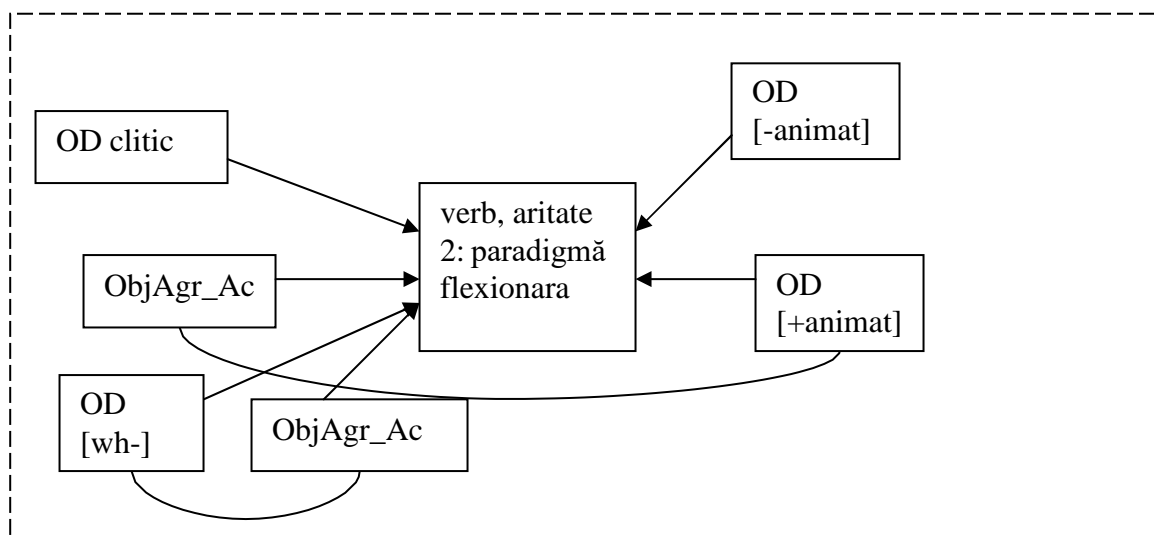


Figura 5.3: Schema generală de reprezentare a grupului verbal cu nucleu bivalent

În majoritatea situațiilor, cliticele apar înaintea verbului. Când condiția fonetică de cliticizare este satisfăcută numai la dreapta, cliticizarea se produce pe auxiliar sau pe verbul lexical. Cliticizarea pe auxiliar este obligatorie (**l a văzut > l-a văzut, *i a spus > i-a spus, *l aș întreba > l-aș întreba, *i aș spune > i-aș spune*). Cliticizarea pe verbul lexical este facultativă, marcând, în esență, o diferență de registru (formal / informal).

Când condiția fonetică de cliticizare este satisfăcută numai la stânga, cliticizarea este facultativă, corelată, din nou, cu diferențe de registru, iar când este realizabilă atât la dreapta, cât și la stânga, cliticizarea se produce obligatoriu pe auxiliar (*nu i-a plăcut, așa i-aș spune*).

Structurile care au clitice aflate după verb sunt specifice unor forme verbale particulare: imperativ, formă afirmativă (*spune-i, pune-l*), gerunziu (*spunându-i, auzind-o*), condițional – optativ și indicativ perfect compus pentru cliticul de acuzativ feminin *o*.

5.3.1 Detalierea structurilor verbale

În realizarea grupului verbal cu nucleu bivalent, un tipar posibil este cel în care complementul direct este exprimat substantival prin acuzativul neprepozițional. În cele două haiku-uri care urmează⁵³ (Ex. 2) rezultatul analizei afișează tiparul verbal cu nucleu bivalent și OD nominal, și arborii sintactici creați plecând de la acest tipar (Fig. 5.4a, Fig. 5.4b):

Exemplul 2:

<i>Privighetoarea</i>	<i>The nightingale</i>
<i>a urmat conservatorul,</i>	<i>attended the Music Academy,</i>
<i>în cuibul matern.</i>	<i>within the maternal nest.</i>

Tiparul verbal:

[V+ NP_{AC}⁵⁴[OD]]ro ☒ [V+ NP_{AC}[OD]]en

Arbori sintactici creați:

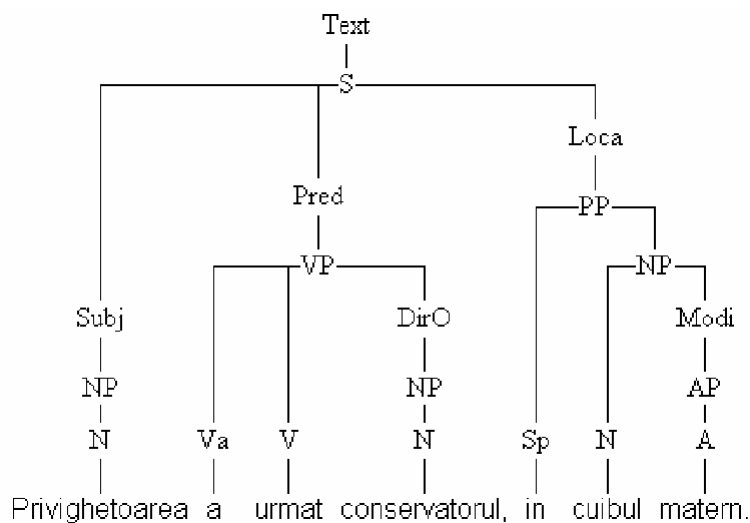


Figura 5.4.a

⁵³ Florin Vasiliu, op. cit., p. 20

⁵⁴ NP_{AC} este notația pe care am folosit-o pentru grupul nominal cu nucleu în cazul acuzativ

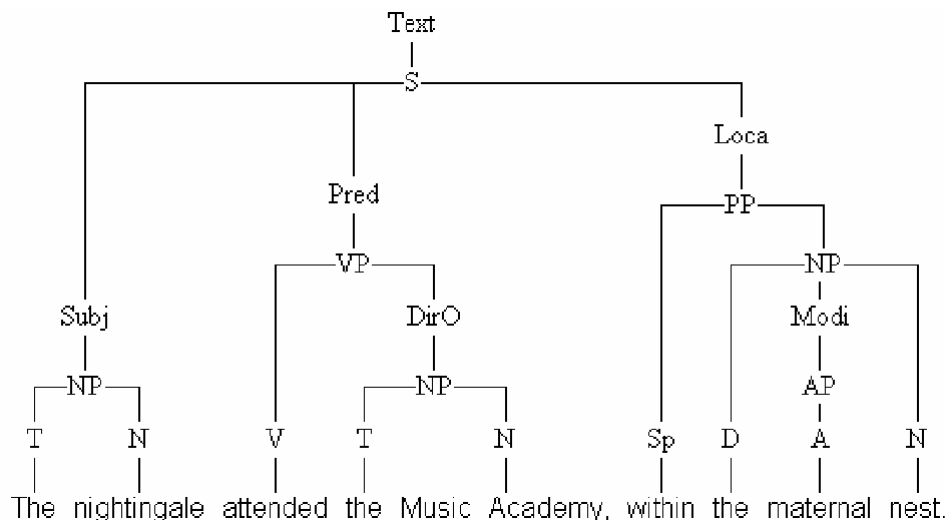


Figura 5.4.b

Complementul direct poate fi exprimat prin două componente, raportate la același verb, dintre care unul este referențial, iar celălalt anaforic: un nominal în acuzativ cu sau fără prepoziția *pe* și un pronume personal sau reflexiv neaccentuat în acuzativ.

Cazul special al anticipării sau reluării obiectului direct prin formele pronominale neaccentuate este rezolvat prin introducerea unei etichete gramaticale speciale, *ObjAgr* (*Acord cu Obiectul*) și prin modificarea regulii de expansiune a grupului nominal [20]. Un exemplu pentru reluarea obiectului direct este analizat în următorul haiku⁵⁵ (Ex. 3), pe baza construirii celor doi arbori sintactici valabili pentru versiunile în limbile română și engleză (Figurile 5.5.a și b):

Exemplul 3:

Cartea Genezei
o citesti in brazdele
negre de april.

The Genesis Book:
you can read it in the black
furrows of April.

⁵⁵ Șerban Codrin, **Dincolo de tăcere**, Editura Haiku, București, 1994, p. 14

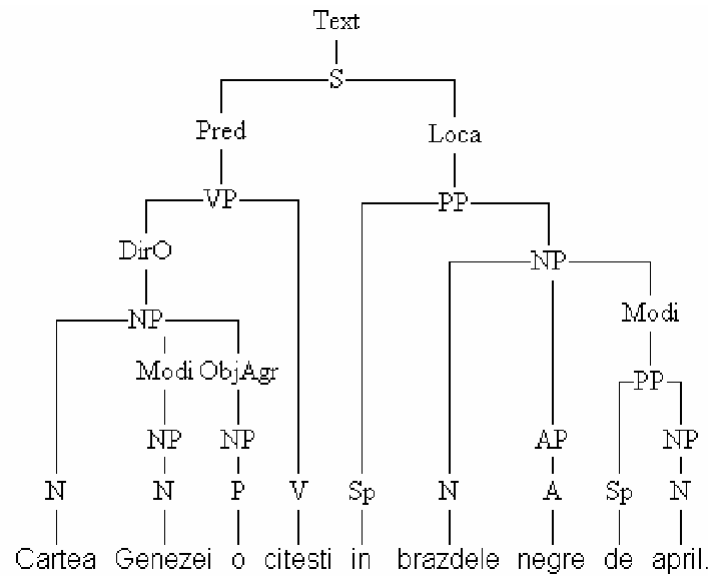


Figura 5.5.a

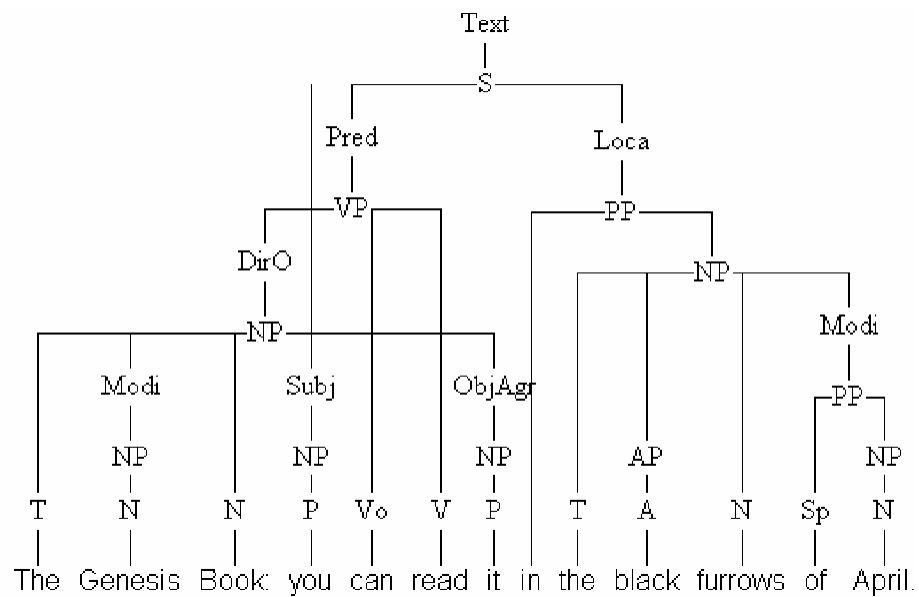


Figura 5.5.b

Tiparele verbale ce rezultă din cele două propoziții arată necesitatea verbului tranzitiv bivalent de a avea o singură dependență asigurată de un complement direct. Trăsătura gramaticală ObjAgr grupează centrul nominal și pronumele nearticulat care a

reluat grupul nominal obiect direct, creând un acord de persoană, număr și caz între substantivul OD și cliticul pronominal⁵⁶:

Tipar verbal:

$[NP_{AC}[N + NP_{ObjAgr}] + V]_{ro} \boxtimes [V + NP]_{en}$

În exemplul care urmează (Ex. 4), pronumele neaccentuat în acuzativ anticipează complementul direct realizat prepozițional în limba română. Pentru limba engleză, situația este mult mai clară, întrucât obiectul direct este realizat referențial:

Exemplul 4:

L-aud venind pe tata, iartă-mă;
*I hear my father coming. Pardon me:*⁵⁷

Tipul de relație sintactică specificat de Exemplul 4 este nondependența, care semnifică termenul echivalent al relației facultative, din teoria lui Hjelmslev⁵⁸. „Este tipul de relație sintactică în care oricare dintre termeni poate fi substituit cu zero. Înglobează relația de coordonare, pe cea de echivalență sau apozitivă, precum și relația care se stabilește între cei doi termeni ai construcțiilor sintactice cu dublare, deci între complement și forma pronominală atonă, “acoperind” situații total diferite ca semantică, dar cu aceeași trăsătură formală a naturii lor facultative”⁵⁹.

În textul din Exemplul 4, termenul substituibil cu zero este cliticul pronominal – dublură a unui complement direct: *L-aud [venind] pe tata < aud [venind] pe tata*.

Descrierea sintactică de tip arborescent include intercalarea gerunziului în limba română, dubla realizare a obiectului direct (referențial și anaforic) și trăsătura gramaticală specială ObjAgr, care guvernează relația dintre nucleul nominal și cliticul pronominal (Figura 5.6 a):

⁵⁶ Întregul compozit $N + NP_{ObjAgr}$ realizează sintactico-semantic obiectul direct.

⁵⁷ William Shakespeare, **Regele Lear**, Ediție bilingvă, Institutul European, 2000, pp. 108- 109

⁵⁸ Louis Hjelmslev, Preliminarii la o teorie a limbii, trad. rom. D. Copceag, București, 1967.

⁵⁹ Valeria Guțu Romalo, Gramatica limbii române, EDP, București, 1973, p. 41

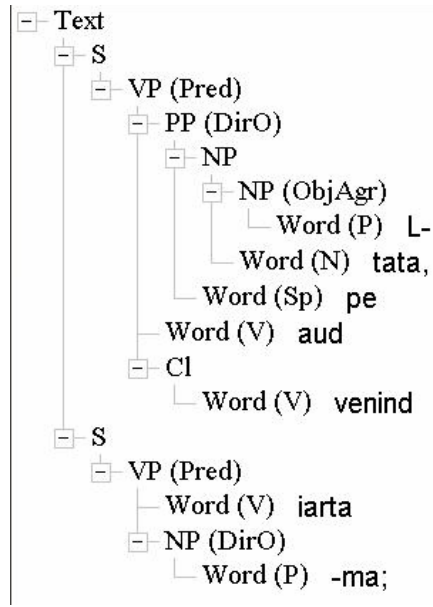


Figura 5.6 a

În limba engleză, descrierea sintactică diferă de cea corespondentă în limba română prin prezența determinantului posesiv, fără însă a mai problematiza intercalarea verbului la modul gerunziu (Figura 5.6 b):

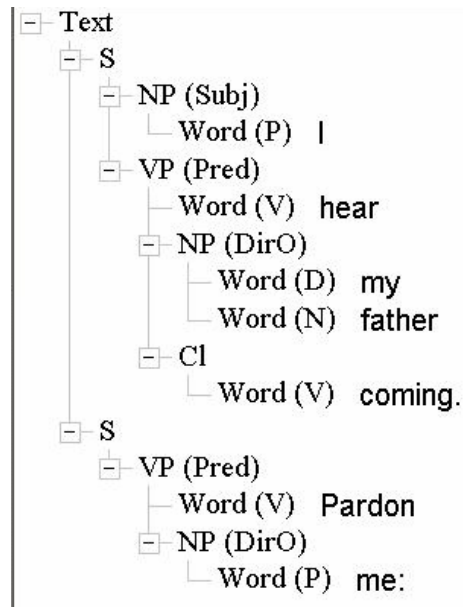


Figura 5.6 b

A doua propoziție din exemplul analizat mai sus dezvoltă un grup verbal format din nucleu la modul imperativ și obiect direct exprimat prin pronume personal în cazul acuzativ. Din cele două propoziții pot fi generate două tipare verbale cu nucleu tranzitiv:

Tipare verbale:

- a. $[NP_{ObjAgr} + V + PP[pe [NP_{AC}]]]_{ro} \boxtimes [V + NP_{AC}]_{en} =$
 $[PP[pe [NP_{AC} [NP_{ObjAgr} + N]]] + V]_{ro} \boxtimes [V +$
 $NP_{AC}]$
- b. $[V + NP_{AC}]_{ro} \boxtimes [V + NP_{AC}]_{en}$

Grupul prepozițional cu funcție de obiect direct, reluat în cadrul grupului verbal prin pronumele clitic de persoana a III-a, formă de acuzativ evidențiază trăsătura gramaticală ObjAgr, stabilind o legătură sintactico-semantică între nucleul nominal și substitutul pronominal care îl reia (Ex. 5, Fig. 5.7 a, b):

Exemplul 5:

pe un nebun nu-l va crede.
*he will not believe a fool.*⁶⁰

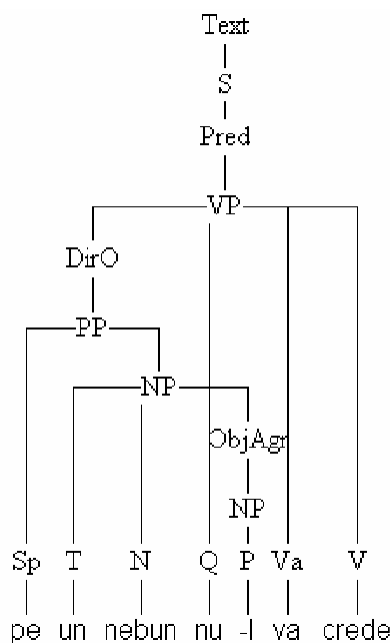


Figura 5.7 a

⁶⁰ W. Shakespeare, op.cit, pp. 86-87

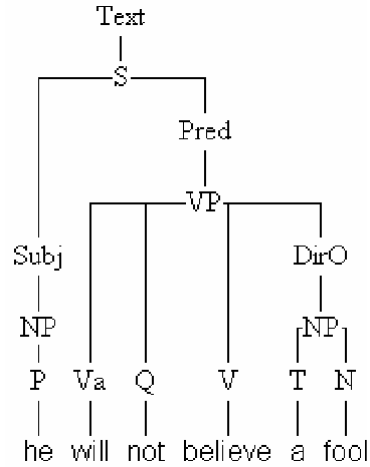


Figura 5.7 b

Tipar verbal:

[PP[pe [NP_{AC} [N + NP_{ObjAgr}]]] + V]_{ro} ☒ [V + NP_{AC}]

Din punctul de vedere al realizării relației de dependență, grupul verbal din Exemplul 6 se bazează pe *dependențe obligatorii*⁶¹, în care determinantul nu poate fi anulat. Este vorba de relația dintre verb și complementul direct: *văd* ← *zeii*, *îi* → *iau*, *îi* → *înșurubez*; *see* ← *the gods*, *take* ← *them*, *thrust* ← *them*:

Exemplul 6:

aieva văd zeii de fildeș,
 îi iau în mână și
 îi înșurubez râzând, în lună,⁶²

I clearly see the ivory gods,
 I take them in my hands and
 thrust them, laughing, in the moon⁶³

⁶¹ În structurile de bază, dependențele obligatorii sunt relațiile dintre diverși determinanți obligatorii și regentul lor, de pildă între subiect și predicat, între numele predicativ și verbul copulativ, între anumite complemente neanulabile: directe (El trimite un mesaj. <*El trimite.), indirecte în dativ (Îi dă mâncare copilului <*Dă mâncare.), indirecte prepoziționale (A apelat la tine <*A apelat.), circumstanțiale (Se comportă civilizată <* Se comportă).

⁶² http://www.imaginelife.ro/poezii/poezii-de-dragoste/nichita_stanescu_-_varsta_de_aur_a_dragostei/

⁶³ http://www.romanianvoice.com/poezii/poezii_tr/goldenage.php

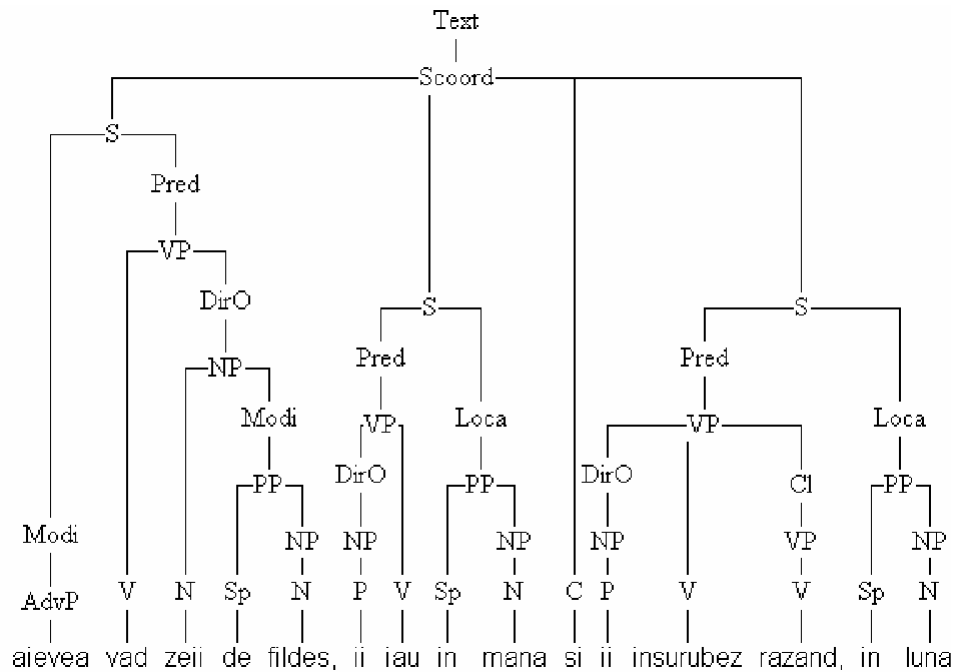


Figura 5.8 a

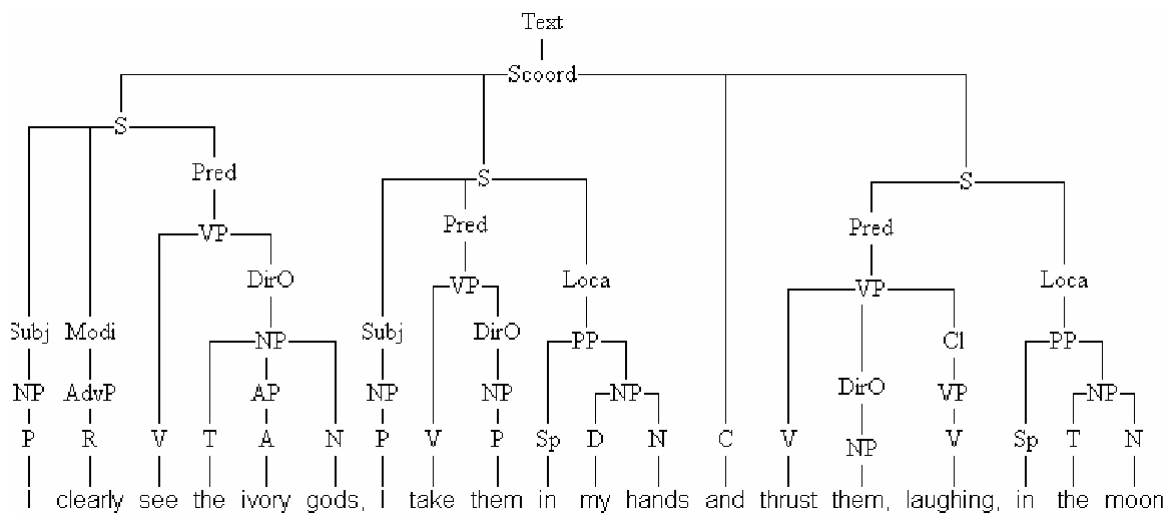


Figura 5.8 b

Tipar verbal:

[V + NP_{AC}] [NP_{AC}[N*⁶⁴]+ V] [NP_{AC}[N*]+ V]_{ro} ☒ [V + NP_{AC}] [V + NP_{AC}[N*]_{en}]
 [V + NP_{AC}[N*]]_{en}

⁶⁴ N* este notația folosită pentru substitutul cu rol de nucleu al grupului nominal, respectiv categoria pronumelui introdusă în biblioteca programului LingTreeConstructor

Obligativitatea determinantului poate avea și cauze semantice, pe lângă cele sintactice, determinantul actualizând un anumit sens al verbului (polisemantic). În exemplul 7, primul sens al verbului *a tăia* este cel de “A despărți, a separa ceva în bucăți cu ajutorul unui obiect tăios sau prin diferite procedee fizice și chimice; a diviza, a scinda, a despica, a fragmenta, a îmbucătăți.”⁶⁵ Sensul actualizat în contextul versurilor stănesciene este însă cel de *a străbate, a traversa*. Izolat de context, verbul este semantic ambiguu; contextului îi revine rolul de dezambiguizare, ceea ce atrage după sine calitatea argumentului *câmpul* ca determinant obligatoriu.

Exemplul 7:

Câmpul tăindu -l, pe două potcoave
calul meu saltă din lut, fumegând⁶⁶

Cutting through the field -up on two shoes
my horse leaps, steaming, from the clay.⁶⁷

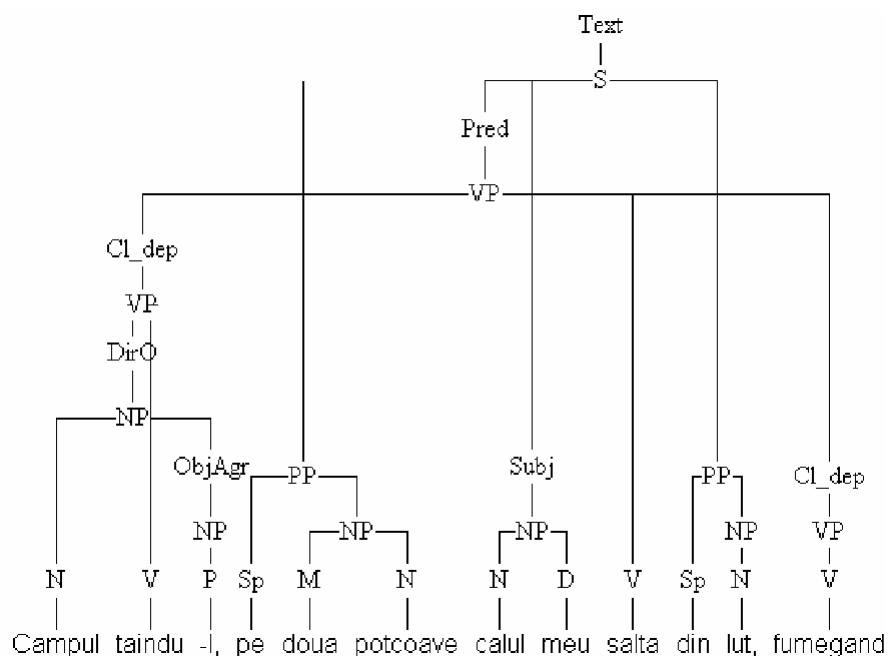


Figura 5. 9 a

⁶⁵ <http://dexonline.ro/definitie/tăia>

⁶⁶ http://art-zone.ro/poezii/nichita_stanescu/O_calarire_in_zori.html

⁶⁷ http://www.romanianvoice.com/poezii/poezii_tr/horseback.php

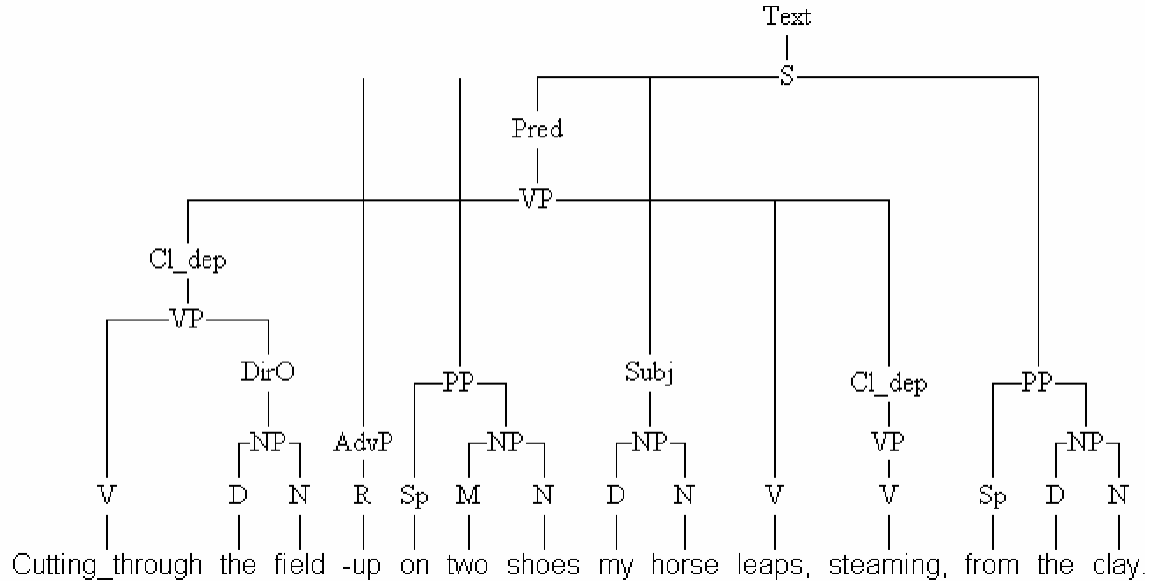


Figura 5. 9 b

Tipar verbal:

$$[\text{NP}_{\text{AC}} + \text{V} + \text{NP}_{\text{ObjAgr}}]_{\text{ro}} \boxtimes [\text{V} + \text{NP}_{\text{AC}}]_{\text{en}} =$$

$$[\text{NP}_{\text{AC}}[\text{N} + \text{NP}_{\text{ObjAgr}}] + \text{V}]_{\text{ro}} \boxtimes [\text{V} + \text{NP}_{\text{AC}}]_{\text{en}}$$

Cele două realizări distincte ale grupului verbal cu nucleu bivalent pot fi urmărite în textul eminescian următor, unde primul grup verbal este format din verb și obiect direct reluat clitic, iar al doilea grup verbal este constituit din verb și obiect direct exprimat referențial:

Exemplul 8:

Lacul codrilor albastru
 Nuferi galbeni îl încarca;
 Tresarind în cercuri albe
 El cutremura o barca.⁶⁸

Water lilies load all over
 The blue lake amid the woods,
 That imparts, while in white circles
 Startling, to a boat its moods.⁶⁹

⁶⁸ http://www.e-referate.ro/referate/Mihai_eminescu_-_lacul2005-03-18.html

Tipar verbal:

[NP_{AC} [N + NP_{ObjAgr}] + V] [V + NP_{AC} [OD]]_{ro} ☒ [V + NP_{AC} [OD]] [V + PP [to [NP_D]] + NP_{AC} [OD]]_{en}

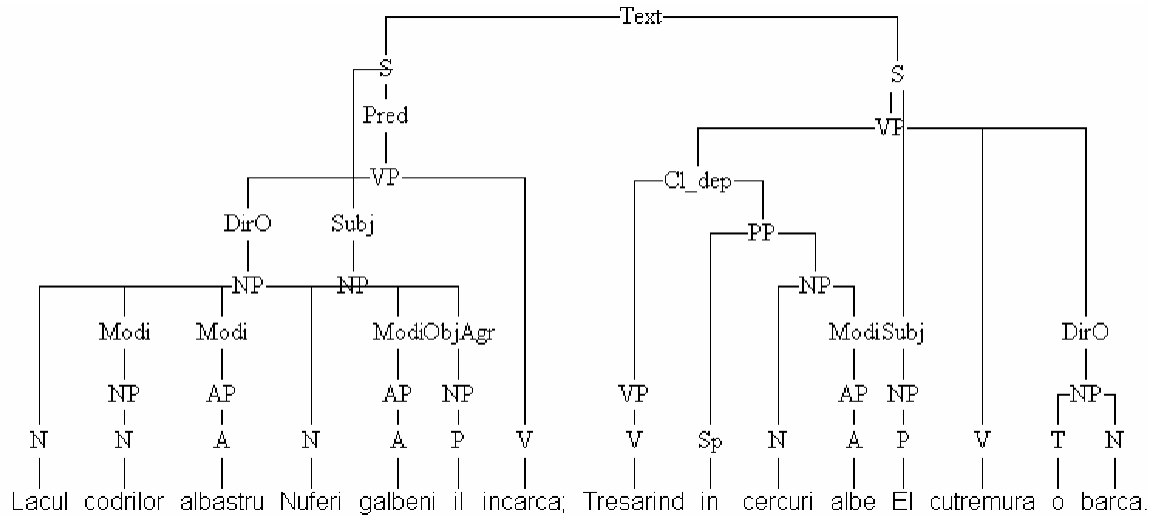


Figura 5.10 a

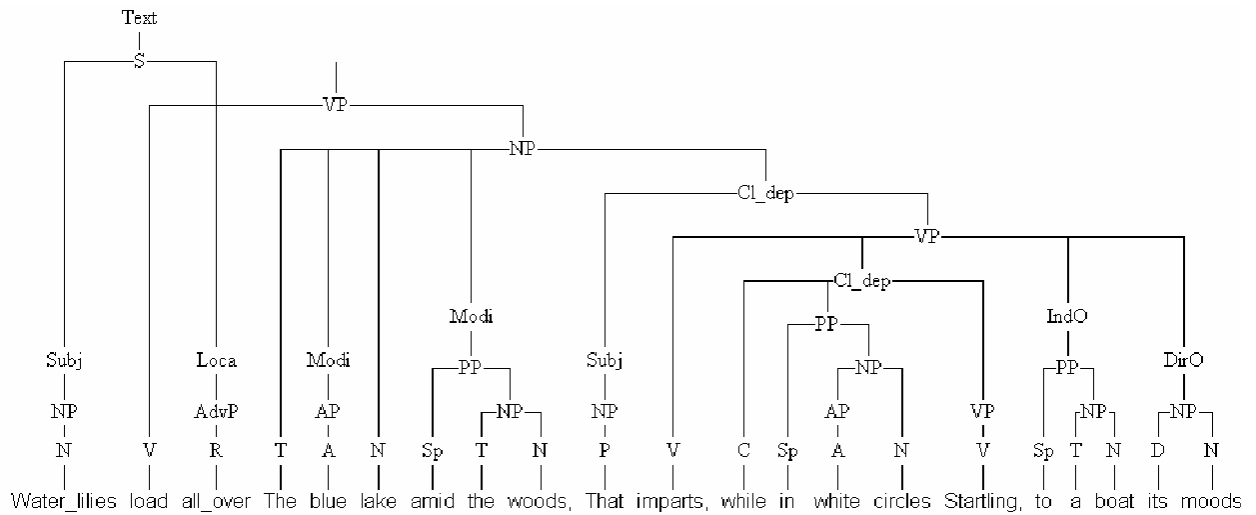


Figura 5.10 b

O intervenție pe care am avut-o asupra bibliotecii programului LingTreeConstructor a avut în vedere editarea fișierului cu tipurile de grupuri sintactice

⁶⁹ http://www.romanianvoice.com/poezii/poezii_tr/lake.php

prin introducerea grupului nominal multiplu, creat prin coordonare: \pt NP; \pt NPcoord⁷⁰:

Tiparul verbal pentru limba română:
[V+ NPcoord [NPAC [OD] și NPAC [OD]]]

Exemplul 9⁷¹:

Au ochi să te-oprească,
Și-aceiași ochi să te gonească
Curând...

Their eyes will seduce you
And the same eyes will chase you
Away...

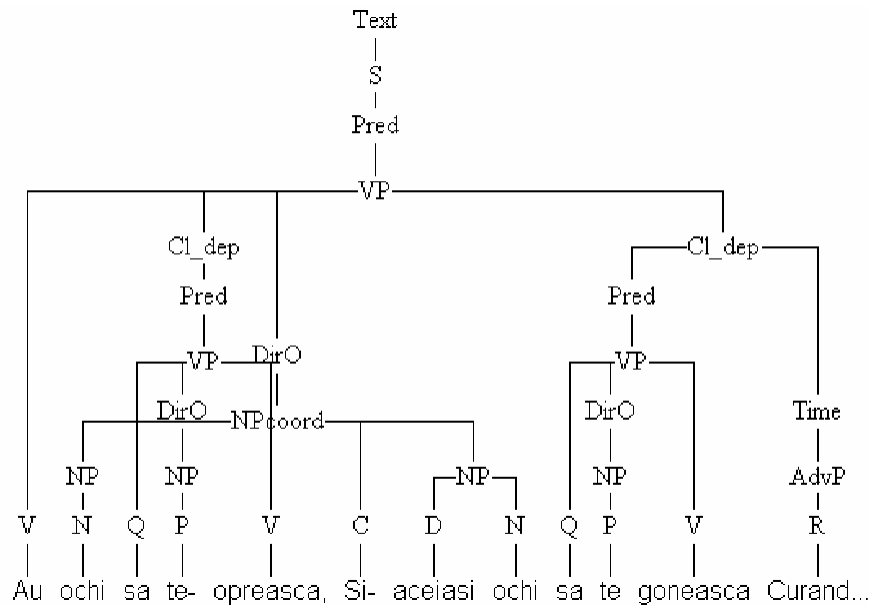


Figura 5.11 a

⁷⁰ \pt este notația folosită pentru tipul de grup sintactic (*phrase type*, în limba engleză).

⁷¹ http://www.fantasypieces.net/translation/alexandru_andries/

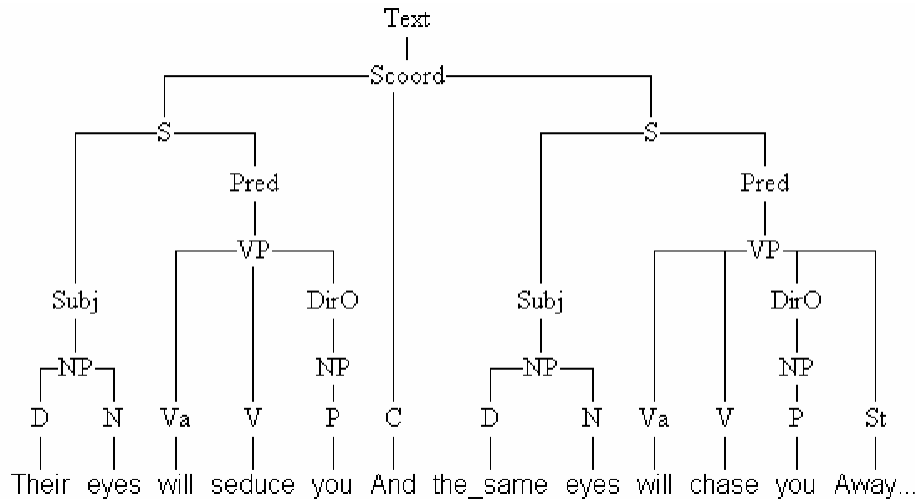


Figura 5.11.b

5.4 Etapa de cercetare. Verbele trivalente

Verbele trivalente sunt caracterizate prin capacitatea combinării cu trei determinări actanțiale obligatorii, fiecare primind un rol tematic diferit și ocupând poziții sintactice diverse: „El **mă** întreabă **ceva**.”; „Ea **îmi** cere **un album**.”

Argumentul-obiect este specificat prin obiect direct, obiect indirect, complement secundar, complement predicativ al obiectului, obiect prepozițional.

5.4.1 Obiectul direct și obiectul indirect

Pe lângă actantul obiect direct structura actanțială a verbelor tranzitive implică alți actanți, printre care cel mai frecvent este obiectul indirect, exprimat sintactic prin complementul indirect.

Complementul indirect este partea secundară de propoziție care determină un verb tranzitiv și, din punct de vedere semantic, indică obiectul asupra căruia se resfrânge indirect acțiunea verbului sau căruia i se atribuie o acțiune, o însușire sau o caracteristică exprimată prin cuvântul determinat.

În funcție de dispunerea complementelor direct și indirect, urmărind, de asemenea, anticiparea sau reluarea unuia din cele două obiecte, au rezultat șapte tipare verbale cu nucleu tranzitiv, pentru diateza activă:

- b.1. (OI clitic + OD clitic) + V
- b.2. V + (OD + OI)
- b.3. ObjAgr_D + V + (OD + OI)
- b.4. ObjAgr_D + V + (OI + OD)
- b.5. V* + (OI + OD)
- b.6. ObjAgr_Ac + V + (OD + OI)
- b.7. ObjAgr_Ac + V + (OI + OD)

Schema obținută prin relaționarea celor șapte tipare este prezentată în figura următoare:

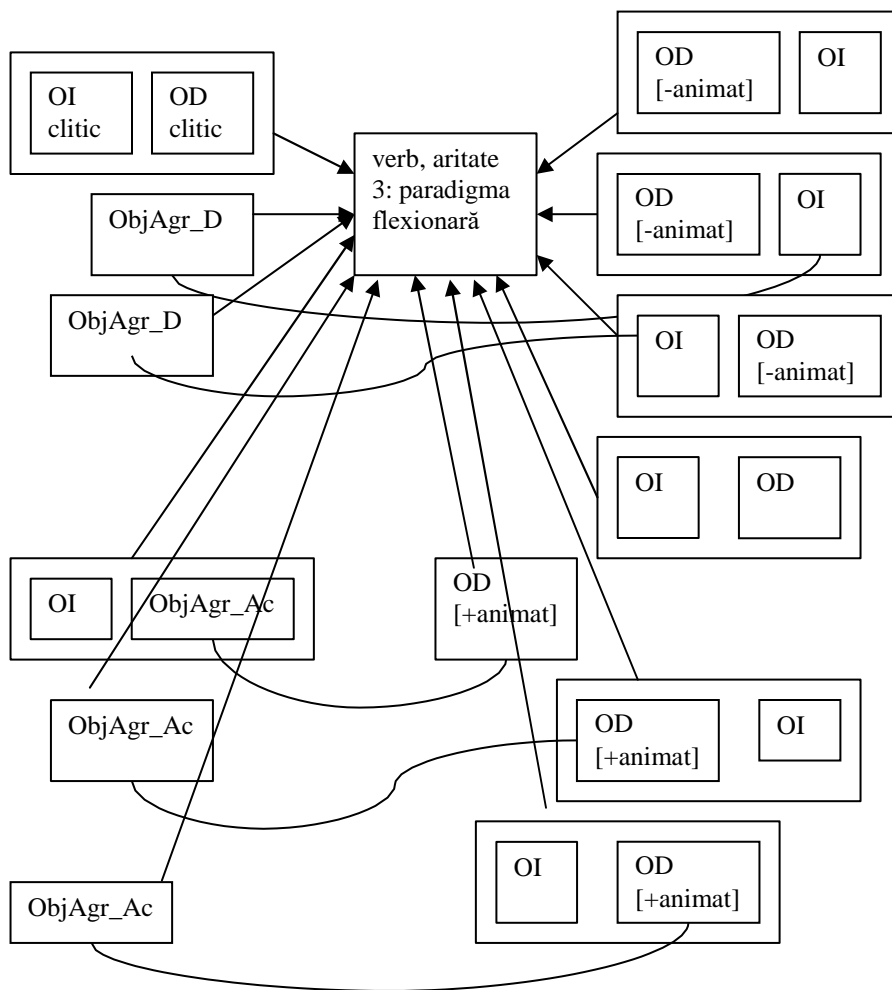


Figura 5.12: Structura verbală cu aritate 3 de tipul (verb tranzitiv, obiect direct și obiect indirect)

5.4.1.1 Dezvoltarea structurilor verbale cu obiect direct și obiect indirect

Structura ierarhică a grupului verbal cu nucleu trivalent este reprezentată grafic în Figura 5.13 a, pentru limba română, și Figura 5.13 b, pentru limba engleză, plecând de la textele a două haiku-uri⁷² (Ex. 10):

Exemplul 10:

<i>Din prag un greier</i>	<i>From the threshold, a cricket</i>
<i>îmi dictează un haiku</i>	<i>orders me a haiku</i>
<i>în formă fixă.</i>	<i>in fixed form.</i>

La nivel descriptiv, scopul este de a surprinde mulțimea regulilor și a restricțiilor care generează exclusiv grupuri verbale gramaticale pentru limbile română și engleză. La nivel explicativ, cei doi arbori sintactici acoperă diferențele superficiale dintre cele două limbi și descoperă principiile de bază care sunt comune limbilor română și engleză.

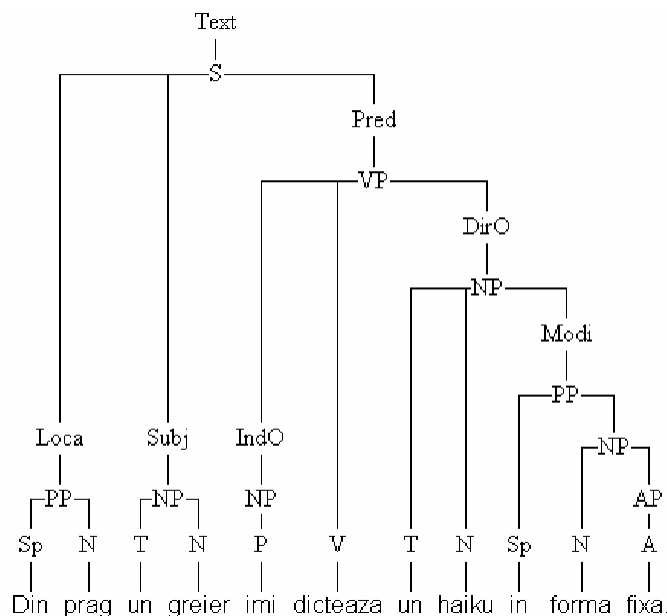


Figura 5.13 a: Deschiderea arborelui sintactic pe orizontală

⁷² Florin Vasiliu, *Tolba cu licurici*, Editura Haiku, București, 1993, p. 67

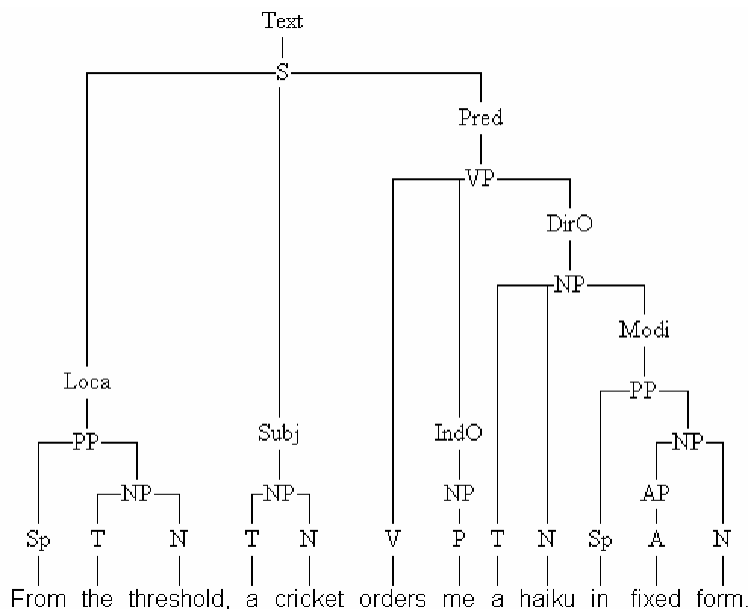


Figura 5.13 b: Deschiderea arborelui pe orizontală

Regulile gramaticale generează, pentru limba română, grupul verbal cu nucleu trivalent păstrând topica IndO+ V+ DirO, iar pentru limba engleză topica se modifică de forma V+ IndO+ DirO. În ambele situații, complementul indirect este clitic, iar complementul direct este exprimat prin substantiv comun. Tiparul verbal cu nucleu tranzitiv, pentru diateza activă este în consecință următorul:

$$[\text{NP}_D + \text{V} + \text{NP}_{AC}]_{ro} \boxtimes [\text{V} + \text{NP}_D + \text{NP}_{AC}]_{en}.$$

În propozițiile echivalente din Exemplul 11 putem analiza topica Obiect indirect + V + Obiect direct, în limba română, respectiv V + Obiect direct + Obiect indirect, în limba engleză:

Exemplul 11:

la ei acasă / le-am dat scrisoarea Înălțimii Voastre
*at their home / I did commend your Highness' letters to them*⁷³

Tipar verbal:

$$[\text{NP}_D + \text{V} + \text{NP}_{AC}]_{ro} \boxtimes [\text{V} + \text{NP}_{AC} + \text{PP}[\text{to} [\text{NP}_D]]]_{en}$$

⁷³ W. Shakespeare, op. cit., pp. 130- 131

Structurile arborescente ordonează în mod diferit cele două complemente cerute de verbul tranzitiv *a da*, observându-se că, pentru limba engleză, grupul nominal cu nucleu clitic în dativ, din română, se transformă în grup prepozițional, introdus prin prepoziția *to* (Figurile 5.14 a și b):

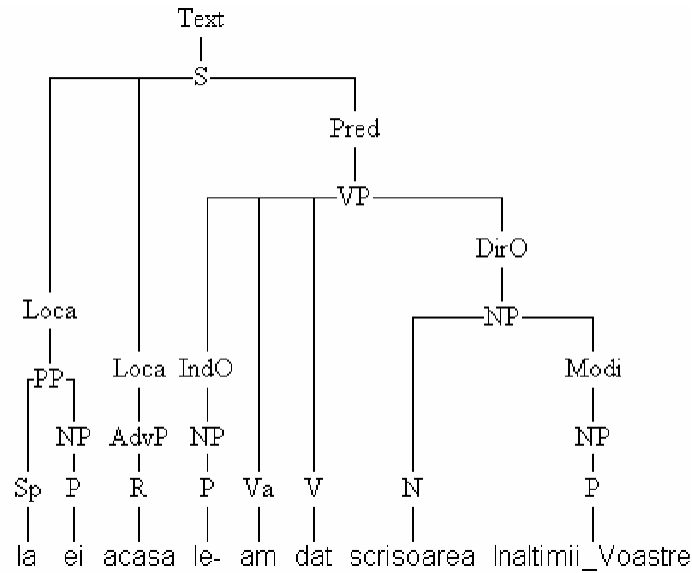


Figura 5.14 a

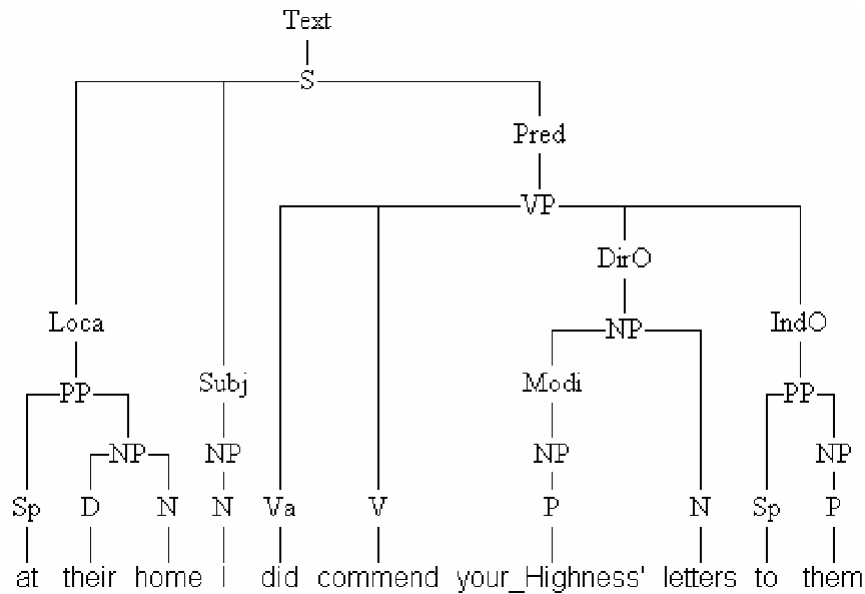


Figura 5.14 b

În propozițiile următoare (Ex. 12), grupul predicatului este format dintr-un verb la modul imperativ, însoțit de două complemente. Atât în limba română (Fig. 5.15 a), cât și în limba engleză (Fig. 5.15 b), ordinea complementelor cerute de verb este următoarea: obiect indirect, cu nucleu anaforic, obiect direct, cu nucleu referențial.

Exemplul 12:

Dă-mi scrisoarea, domnule.

Give me the letter, sir⁷⁴.

Tipar verbal:

$[V + NP_D + NP_{AC}]_{ro}$ ☒ $[V + NP_D + NP_{AC}]_{en}$

Structurile arborescente arată un grup al predicatului (verb + cele două complemente) și un grup nominal în vocativ:

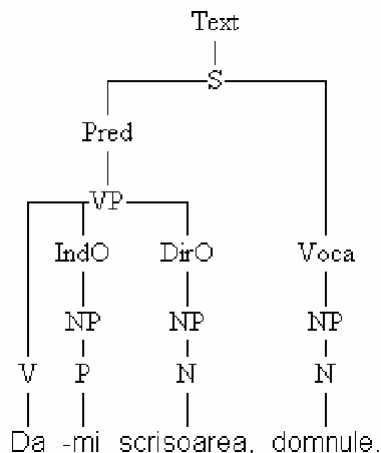


Figura 5.15 a

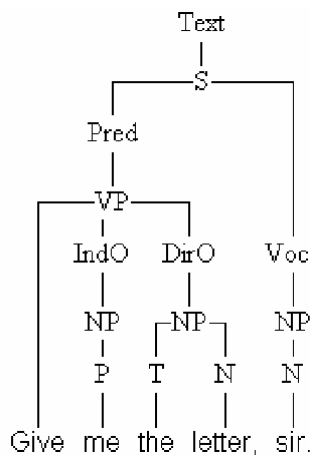


Figura 5.15 b

⁷⁴ W. Shakespeare, op. cit., pp. 68-69

Pronumele reflexiv *ne* din exemplul următor are funcție sintactică de obiect direct⁷⁵, iar pronumele personal în dativ, *lui*, este obiectul indirect, ambele argumente fiind cerute de verbul *a închina*. În limba engleză, traducerea transformă cele două complemente, propunând doar un complement prepozițional:

Exemplul 13:

vor așeza acolo un zeu
să ne-nchinăm lui⁷⁶

they will place a god there
for us to worship⁷⁷

Tipar verbal:

[V+ NP_{AC} [OD]] [NP_{AC} [N*]+ V+ NP_D [N*]]_{ro} ☒ [V+ NP_{AC} [OD] + PP [for [NP_{AC}]]]_{en}

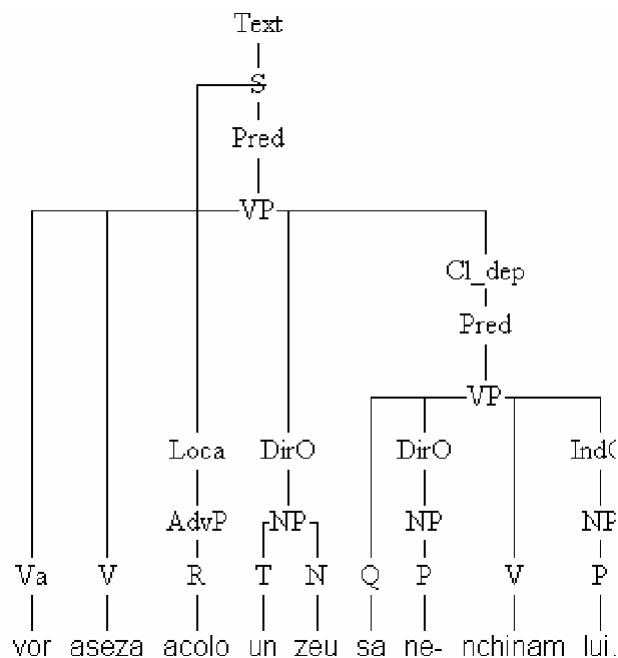


Figura 5.16 a

⁷⁵ Formele neaccentuate ale pronumelui reflexiv au funcție sintactică numai dacă pot fi dublate printr-o formă corespunzătoare persoanei și cazului.

⁷⁶ http://ro.wikisource.org/wiki/Elegia_a_doua_Getica

⁷⁷ http://www.romanianvoice.com/poezii/poezii_tr/elegy2.php

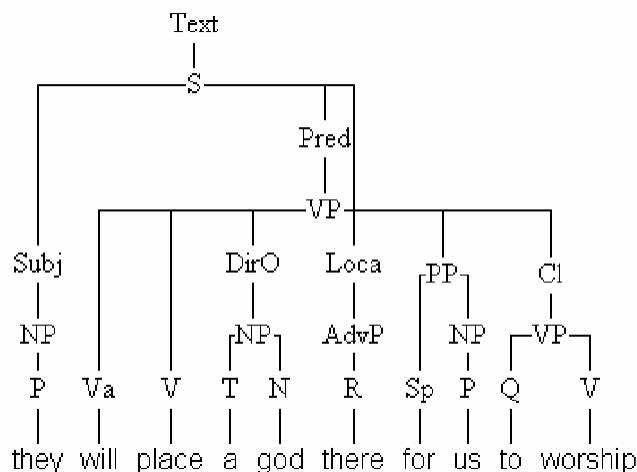


Figura 5.16 b

5.4.2 Structuri verbale cu obiect direct și complement secundar

Un alt actant al verbelor tranzitive este un al doilea obiect direct, numit de G. Pană Dindelegan obiect secundar. Verbele tranzitive care se construiesc cu un obiect secundar se numesc dublu tranzitive. Informația oferită de gramaticile curente cu referire la verbele dublu tranzitive se reduce la constatarea faptului că există un grup restrâns de verbe ce exprimă acțiuni care se extind asupra a două obiecte directe [97]. Aceste verbe pot avea în calitate de determinanți obligatorii două complemente directe: unul al persoanei (cu trăsătura semantică *uman*) și unul al obiectului (cu trăsătura semantică *non-uman*). Drept exemplu se aduc, de obicei, verbele: *a întreba*, *a învăța*, *a ruga*, *a examina* etc. în contexte de tipul: *Profesorul îl întreabă pe elev tema* [15: 134].

Cel de-al doilea complement direct (al obiectului) guvernat de verbele tranzitive se distinge de complementul direct al persoanei printr-o serie de caracteristici, relevate de Gabriela Pană Dindelegan [69: 19-20], astfel:

- este imposibilă substituția acestui complement cu un pronume în acuzativ;
- este imposibilă apariția prepoziției *pe* înaintea complementului direct al obiectului;
- este imposibilă coordonarea acestor două tipuri de complemente în cadrul construcției dublu tranzitive **L-am învățat pe Ion și ceva*;
- este imposibilă dublarea complementului direct al obiectului.

Când verbul tranzitiv contractează legături cu un obiect direct și un complement secundar, rezultă două structuri verbale corespunzătoare arității 3:

1. OD clitic + V + CS
2. ObjAgr_Ac + V + OD + CS

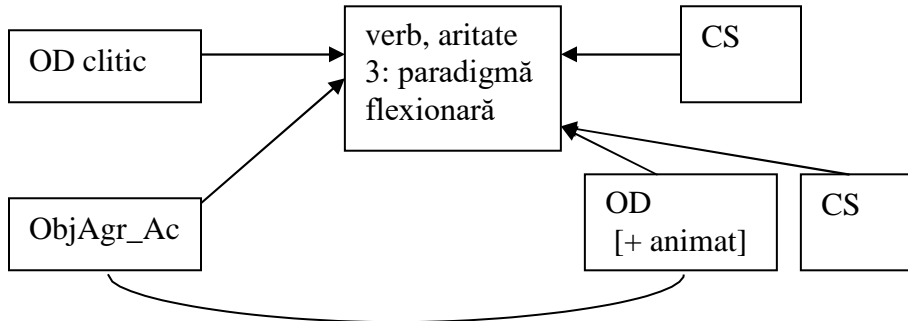


Figura 5.17: Structuri verbale cu obiect direct și complement secundar

5.4.2.1 Dezvoltarea structurilor verbale cu obiect direct și complement secundar

1. OD + V + CS // OD clitic pronominal, CS = complement secundar; pentru V de tipul: {a învăța, a asculta, a întreba, a anunța, a ruga, a sfătui, a îndruma}.

Gramatica Academiei (vol. 2, pp. 392-393) consideră complementul secundar ca fiind un termen subordonat față de centrul grupului verbal, într-o structură ternară din care face parte și un complement direct subordonat aceluiași centru verbal. Cele două componente au roluri tematice distincte, și anume: *Pacient* pentru complementul secundar, *Țintă* pentru complementul direct (*El anunță pe cineva ceva.*); *Locativ* pentru complementul secundar, *Pacient* pentru complementul direct (*El îl trece pe bătrân strada.*), etc.

În structura descrisă mai sus, complementul secundar este realizat morfologic astfel: {substantiv comun [-animat]; pronume [-animat]: demonstrativ *asta, aia,*

interrogativ- relativ *ce*, nehotărât *altceva, ceva, orice*, locuțiunile pronominale *cine știe ce, nu știi ce, te miri ce*, pronume negativ *nimic* }.

2. ObjAgr_Ac + V + OD + CS

De exemplu: „**I**-a învățat **pe copii o colindă.**”, „**Ne**-a rugat **pe noi ceva.**”, „Nu **i**-a rugat **pe aceștia nimic.**”

Propoziția cu verbul la modul imperativ “Întreabă-mă orice” este tradusă în limba engleză prin “Ask me anything” și implică existența unui obiect direct exprimat prin pronumele personal, formă neaccentuată, *-mă*, și a unui complement secundar realizat morfologic prin pronumele nehotărât *orice*. Acest exemplu, analizat, conduce spre arborii sintactici din figura de mai jos:

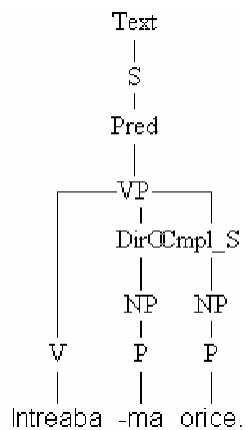


Figura 5.18 a

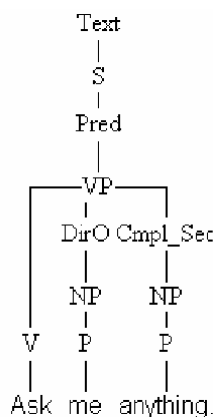


Figura 5. 18 b

Tipar verbal:

$[V + NP_{AC} [OD] + NP_{AC} [Compl_Sec]]_{ro} \boxtimes [V + NP_{AC} [OD] + NP_{AC} [Compl_Sec]]_{en}$

În propozițiile echivalente “Scopul meu este de a te învăța matematică” \boxtimes “My goal is to teach you math”, complementul secundar este exprimat prin substantiv comun, iar cliticul pronominal precede verbul aflat la forma de infinitiv (vezi arborii sintactici din Fig. 5.19 a, b):

Tipar verbal:

$[NP_{AC} [OD] + V + NP_{AC} [Compl_Sec]]_{ro} \boxtimes [V + NP_{AC} [OD] + NP_{AC} [Compl_Sec]]_{en}$

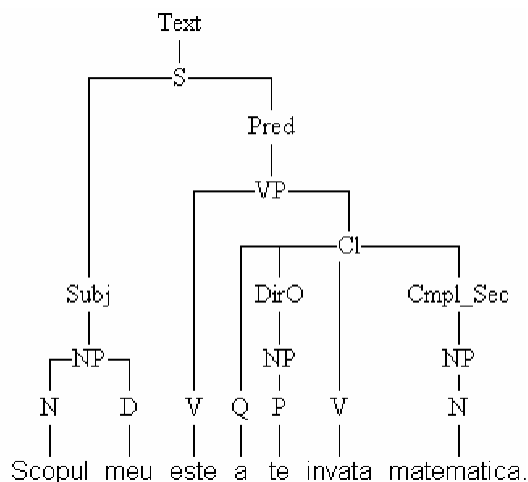


Figura 5.19 a

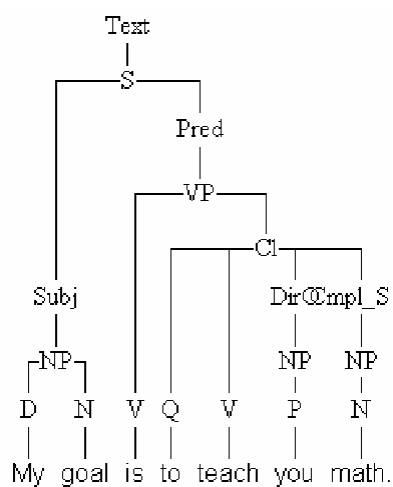


Figura 5.19 b

5.4.3 Structuri verbale cu obiect direct și complementul predicativ al obiectului

Pentru verbe din clasa semantică: *a boteza, a chema, a denumi, a intitula, a numi, a porecli, a alege, a angaja, a desemna*, grupul verbal include ca argumente interne un obiect direct și un complement predicativ al obiectului. Se creează astfel două structuri verbale, în funcție de modul în care obiectul direct este sau nu anticipat printr-o formă pronominală neaccentuată:

1. V + OD clitic + CPO
2. ObjAgr_Ac + V + OD + CPO

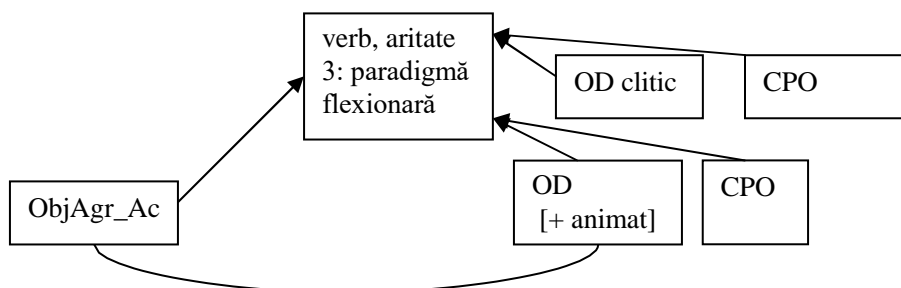


Figura 5.20: Structuri verbale cu obiect direct și complement predicativ al obiectului

5.4.3.1 Dezvoltarea structurilor verbale cu obiect direct și complementul predicativ al obiectului

1. V + OD + CPO // OD este realizat clitic pronominal, CPO este complementul predicativ al obiectului, pentru V: { *a boteza, a chema, a denumi, a intitula, a numi, a porecli, a alege, a angaja, a desemna*};

De exemplu: „Au denumit **proiectul altfel**.”, „Vor alege **plecarea ca soluție**.”, „Partenerii mei vor numi **afacerea drept oportunitate**.”

2. ObjAgr_Ac + V + OD + CPO.

De exemplu: „**L**-au ales **pe tatăl** meu **ca președinte**.”, „**I**-au desemnat **pe aceștia drept înlocuitori**.”

În [1]⁷⁸, complementul predicativ al obiectului este descris ca fiind o clasă de substituție având ca termen prototipic substantivul neprepozițional cu formă de nominativ-acuzativ, iar ca relație sintagmatică specifică, o relație obligatoriu ternară, implicând verbul-centru și un complement.

În crearea arborilor sintactici pentru propozițiile din exemplele următoare, am introdus în LingTreeConstructor eticheta sintactică de complement predicativ al obiectului (Cmpl_Pred_Ob), rezultând, pentru grupul predicatului, cele două tipare:

- a. $[NP_{AC}[OD]^+ V^+ NP_{AC}[Cmpl_Pred_Ob]]_{ro}$  $[V^+ NP_{AC}[Cmpl_Pred_Ob]]_{en}$

Exemplul 14:

It'll make him a big star.⁷⁹

Îl va face o mare stea.

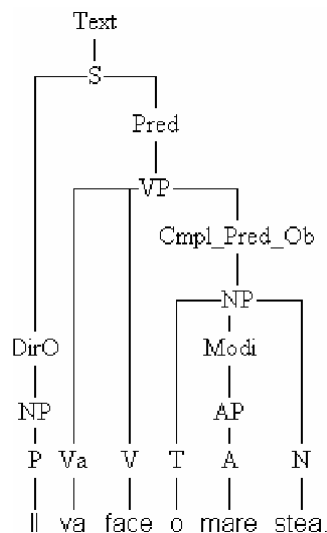


Figura 5. 21 a

⁷⁸ vol. 2, p. 290

⁷⁹ <http://www.imdb.com/title/tt0068646/quotes>

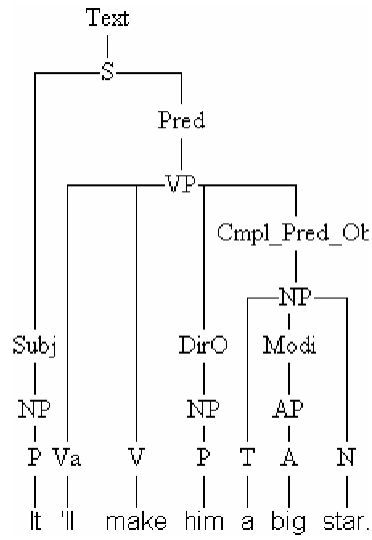


Figura 5. 21 b

Exemplul 15:

Maya îl numește prietenul ei.

Maya calls him her friend.

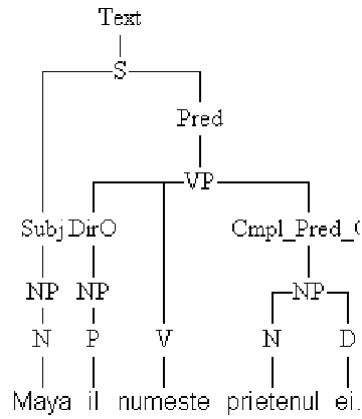


Figura 5. 22 a

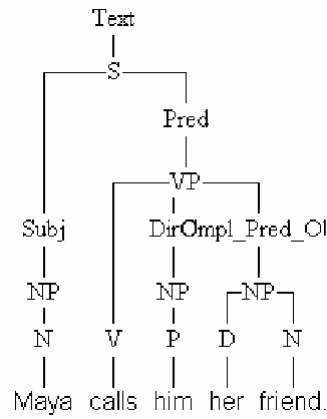


Figura 5. 22 b

b. $[NP_{ObjAgr+} V+ PP [pe [NP_{AC}]_+ PP [Cmpl_Pred_Ob]]_{ro} \boxtimes [V+ NP_{AC+} PP[Cmpl_Pred_Ob]]_{en} =$
 $[PP [pe [NP_{AC}[NP_{ObjAgr+} N]]]_+ V+ PP [Cmpl_Pred_Ob]]_{ro} \boxtimes$
 $[V+ NP_{AC+}$
 $PP[Cmpl_Pred_Ob]]_{en}$

Exemplul 16:

Clubul l-a ales pe Tom ca preşedinte.

The club elected Tom as president.

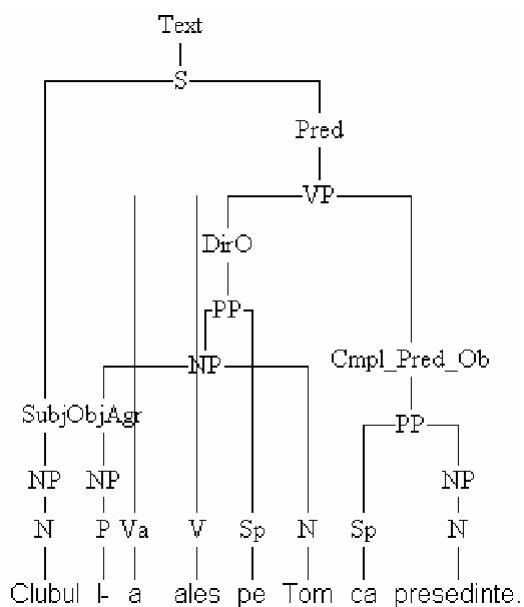


Figura 5. 23 a

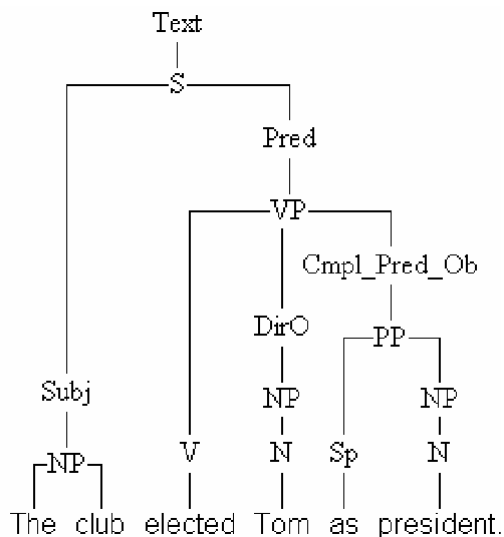


Figura 5. 23 b

În exemplul 16 putem identifica *relația de echivalență*, definită ca un tip de relație sintactică între doi sau mai mulți termeni cu același referent sau coreferențiali și care stabilesc aceeași relație față de un alt termen. Termenii *pe Tom* și *l-* au același referent și sunt subordonați verbului *a ales*.

5.4. Structuri verbale cu obiect direct și obiect prepozițional

Prepoziția și conjuncția reprezintă clasa relatemelor, în sensul că, alături de alte elemente lingvistice (relative pronominale, adjectivale și adverbiale), „[...] asigură desfășurarea relațiilor sintactice – condiție fundamentală a unității structurale a enunțului sintactic” [48].

Disponibilitățile combinatorii ale verbului care guvernează mai multe cazuri în relație cu numele constituie baza unei clasificări sintagmatice a verbelor propuse de V. Guțu Romalo [44] pe care am prezentat-o în introducere. Autoarea omite însă determinanții prepoziționali ai verbului, relație care este urmărită din perspectivă distribuțională de G. Pană Dindelegan în mai multe lucrări [69, 70, 71].

Aplicând testul substituției cu o formă cazuală sau adverbială, G. Pană Dindelegan delimitează clase de verbe care se construiesc cu determinanți prepoziționali cu valoare /±circumstanțială/ (după cum acceptă sau nu substituția cu un adverb din clasa {*unde, cum, când, cât*}).

Verbele cu obiect prepozițional au specificate în partea dreaptă prepozițiile de legătură între nucleul verbal și argumentul – obiect. Aceste prepoziții sunt strict necesare pentru complinirea valenței a treia (de exemplu: *a priva_de*; *a predispune_la*, *a povesti_despre*, etc.).

Structurile verbale cu obiect direct și obiect prepozițional pot avea următoarea dispunere a constituenților:

1. V + OD clitic + OP
2. ObjAgr_Ac + V + OD + OP

5.4.4.1 Dezvoltarea structurilor verbale cu obiect direct și obiect prepozițional

1. V + OD + OP;

De exemplu: „El nu deosebește binele de rău.”, „Au întrebat **copiii despre altceva.**”

2. ObjAgr_Ac + V + OD + OP.

De exemplu: „Muzica **îl** predispune **pe artist la visare.**”, „Bunicii **i-au** îndemnat **pe nepoți la învățătură.**”

Complementul prepozițional este exprimat prin: {substantiv comun; pronume: demonstrativ *asta, aia*, interogativ- relativ *ce, nehotărât *alta, altceva, câteva, ceva, orice, tot, una**, locuțiunile pronominale *cine știe ce, nu știu ce, te miri ce*, pronume negativ *nimic*}.

Structura (complement direct + complement indirect) a fost descrisă separat, în cadrul tiparelor verbale ternare, ca, de altfel, și constructele sintactice complexe, unde OD este anticipat prin forme pronominale personale neaccentuate (vezi trăsătura gramaticală ObjAgr).

Din punct de vedere sintactic, complementul prepozițional funcționează în grupul verbal ca termen subordonat unei relații de dependență față de centrul grupului, care impune nominalului o anumită prepoziție și îi atribuie un anumit rol semantic.⁸⁰

Pe baza criteriului substituției cu o formă causală sau adverbială, se observă că verbul *a se gândi* din Exemplul 17 se combină cu un determinant nesubstituibil, rezultând forma *a se gândi la*. Verbul se construiește cu un obiect prepozițional obligatoriu și se caracterizează prin următoarele particularități:

- prepoziția este univoc determinată, în sensul că verbul *a se gândi* cere în mod necesar prepoziția *la*, și nu alta;
- suprimarea grupului prepozițional nu este posibilă, enunțul **Și într-un ceas gândești* nefiind corect, acceptat, ceea ce arată că determinantul

⁸⁰ Gramatica Academiei, vol. 2, p. 416

prepozițional- obiectul prepozițional- este un constituent obligatoriu, neomisibil.

Exemplul 17:

Și tu citești scrisori din roase plicuri
Și într-un ceas gândești la viața toată.⁸¹

And you're reading old letters, tattered and fading
And retrace a whole life-time in just one hour.⁸²

Tipar verbal:

[V+ NP_{AC} [OD]] [V+ PP [la [NP_{AC}]]]_{ro} ☒ [V+ NP_{AC} [OD]] [V+ NP_{AC} [OD]]_{en}

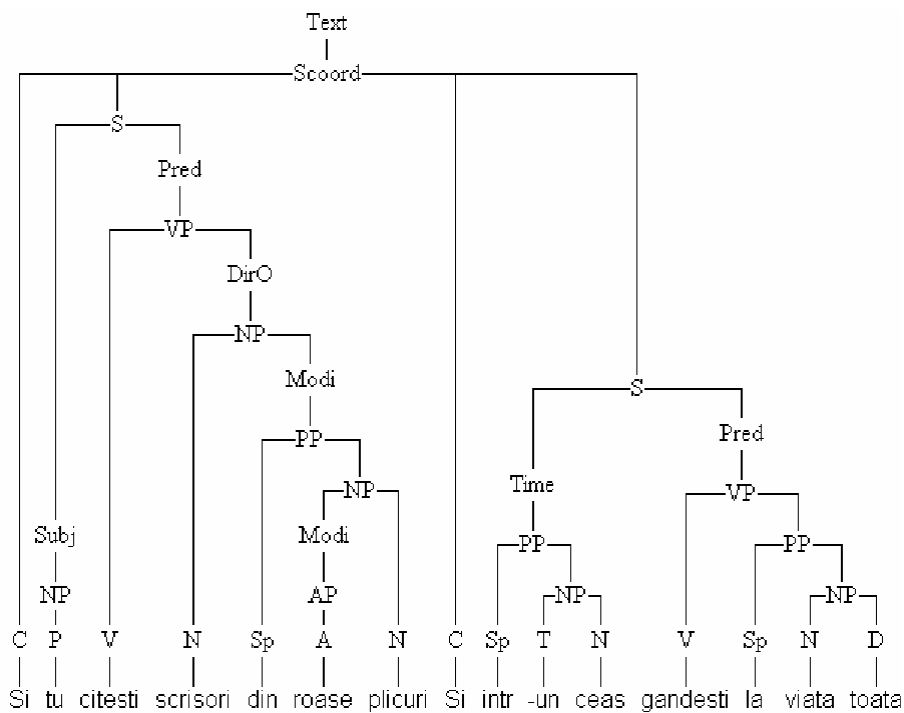


Figura 5.24 a

⁸¹ <http://www.poeziile.com/autori/Mihai-Eminescu/toamna67.php>

⁸² http://www.romanianvoice.com/poezii/poezii_tr/autumn.php

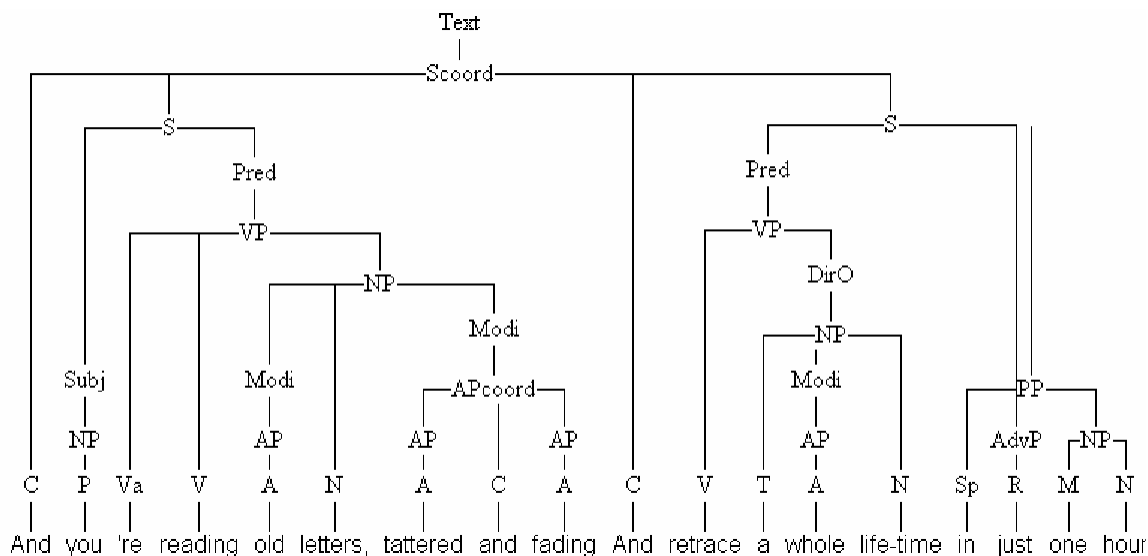


Figura 5.24 b

Pentru construirea celor doi arbori sintactici din Figurile 5. 24 a și 5. 25 b am editat biblioteca LingTreeConstructor asociată fișierului de noduri sintactice, prin crearea nodului de *propoziție coordonată*, definit astfel:

```
\Is Scoord83
\Il Sentence Coord
```

Atunci când în relația [verb + prep + nume] grupul prepozițional este substituibil cu un adverb relativ {*unde, când, cum*}, se consideră că acest grup prepozițional are valoare adverbială și funcție de circumstanțial. G. Pană Dindelegan stabilește și în acest caz două subclase de verbe în funcție de caracterul /±obligatoriu/ (neomisibil) al determinantului prepozițional.

În Exemplul 18, verbul *a înconjura* se construiește cu o anumită prepoziție și cu un circumstanțial de mod obligatoriu: *a înconjura cu*. Obligativitatea determinantului este explicitată de definiția lexicografică, aceasta indicând, între paranteze, selecția în termeni lexicali a circumstanțialului:

⁸³ Notația Scoord apare în arborele sintactic, iar notația Sentence Coord identifică în interfața LingTreeConstructor nodul sintactic pe care îl creăm.

ÎNCONJURĂ vb. 1. v. ocoli. 2. v. ocoli. 3. v. împrejmui. 4. v. încercui. 5. v. asedia. 6. a împresura, a încercui, (fig.) a încinge. (Munții ~ câmpia.) 7. a chenăru, a împrejmui, a încadra, a mărgini, (fig.) a tiv. (Soarele ~ cu lumină poiana.)⁸⁴

Exemplul 18:

Ți-am dat palidele raze ce-nconjoară cu magie
Fruntea îngerului-geniu⁸⁵

I had given you the pale rays which pour magic
On the brow of genius-angels⁸⁶

Tipar verbal:

[NP_D [OI]+ V+ NP_{AC} [OD]] [V+ PP [cu [NP_{AC}]]+ NP_{AC} [OD]]_{ro} ☒ [V+ NP_D [OI]+ NP_C [OD]] [V+ NP_{AC} [OD] + PP[on [NP_{AC}]]]_{en}

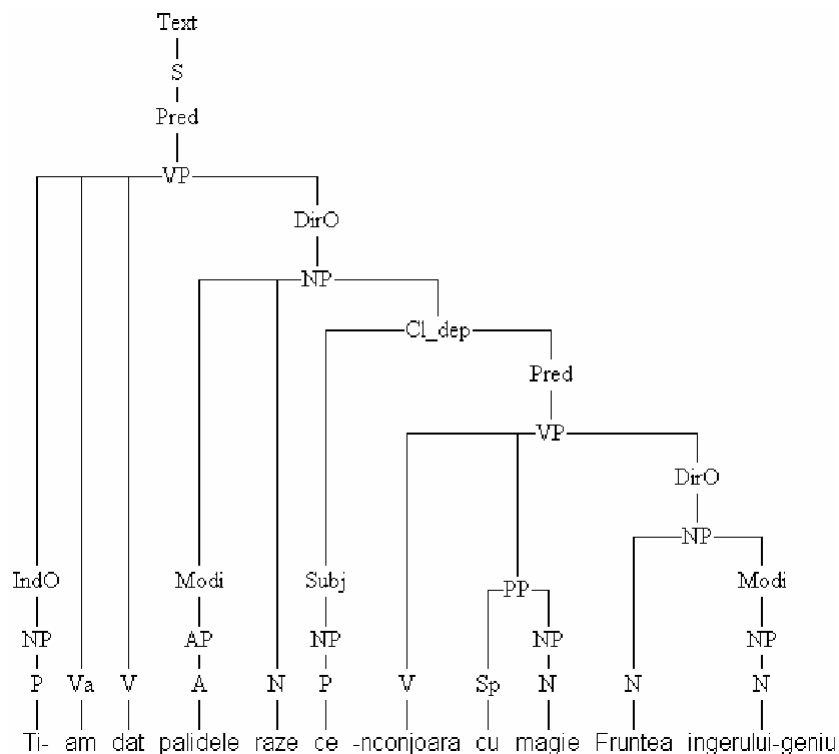


Figura 5. 25 a

⁸⁴ <http://dexonline.ro/definitie/inconjura>

⁸⁵ <http://www.dindragoste.ro/poezii-de-dragoste/mihai-eminescu/venere-si-madona.php>

⁸⁶ http://www.romanianvoice.com/poezii/poezii_tr/venus.php

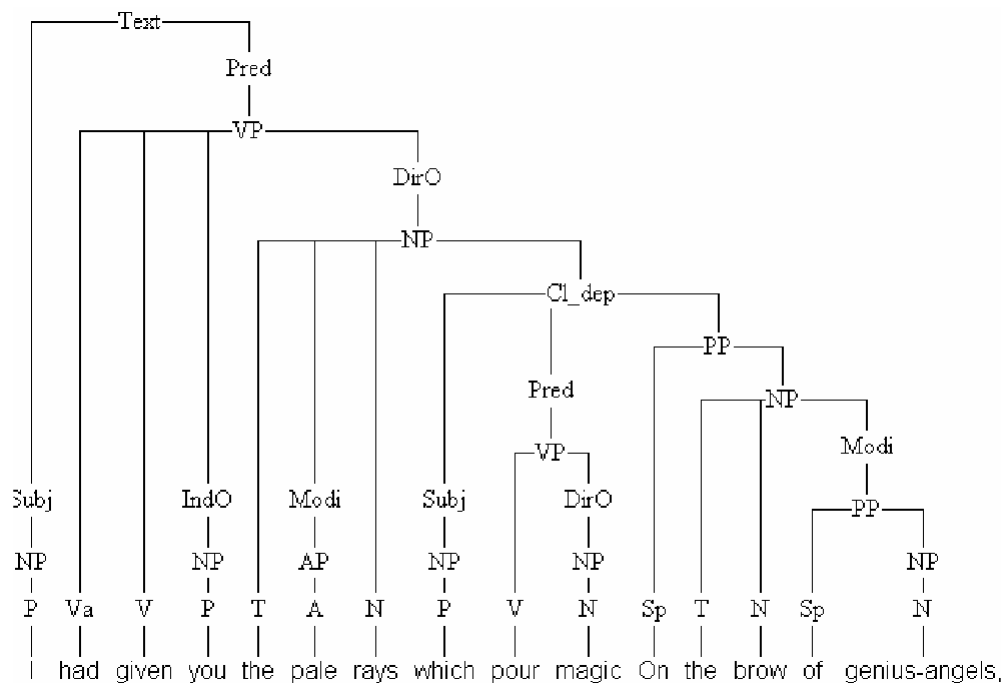


Figura 5. 25 b

Alegerea prepoziției *în* de către verbul *a transforma* din Exemplul 19 este dictată de considerente semantice (*a transforma ceva în altceva*):

A TRANSFORMĂ *transfórm tranz.* 1) *A face să se transforme; a schimba; a preface; a preschimba; a modifica; a converti.* 2) (*valori, funcții, construcții sintactice etc.*) *A echivala cu un alt corespondent, schimbând unele elemente și păstrând altele.* [Sil. *trans-for-*] /<fr. *transformer*, lat. *transformare*⁸⁷

Din perspectiva relației de combinare dintre verb și prepoziție, lexemul *transformă* este un verb construit cu o anumită prepoziție și un circumstanțial obligatoriu. Verbul *a intra*, din același exemplu, se combină cu un circumstanțial facultativ, ceea ce înseamnă că are posibilitatea să aleagă între mai multe prepoziții și chiar termenul \emptyset , de exemplu: *Închid fereastra să nu intre-n casă / Închid fereastra să nu intre (\emptyset):*

⁸⁷ <http://dexonline.ro/definitie/transforma>

Tipar verbal:

[PP [peste [NP_{AC}]]+ V] [V+ NP_{AC} [OD]] [V+ NP_{AC} [OD]+ PP [în [NP_{AC}]]]ro ☒

[V+ NP_{AC} [OD]] [V+ NP_{AC} [OD]] [V+ NP_{AC} [OD]+ PP [into [NP_{AC}]]]en

Exemplul 19⁸⁸:

Zile egale peste mine apasă,
 Închid fereastra să nu intre-n casă.
 Urma lor fină
 Transformă camera mea în vitrină...

Equal days are bearing down on me,
 I have closed my window so they cannot get in.
 Their trace so fine
 Turn my room into a shop window sign...

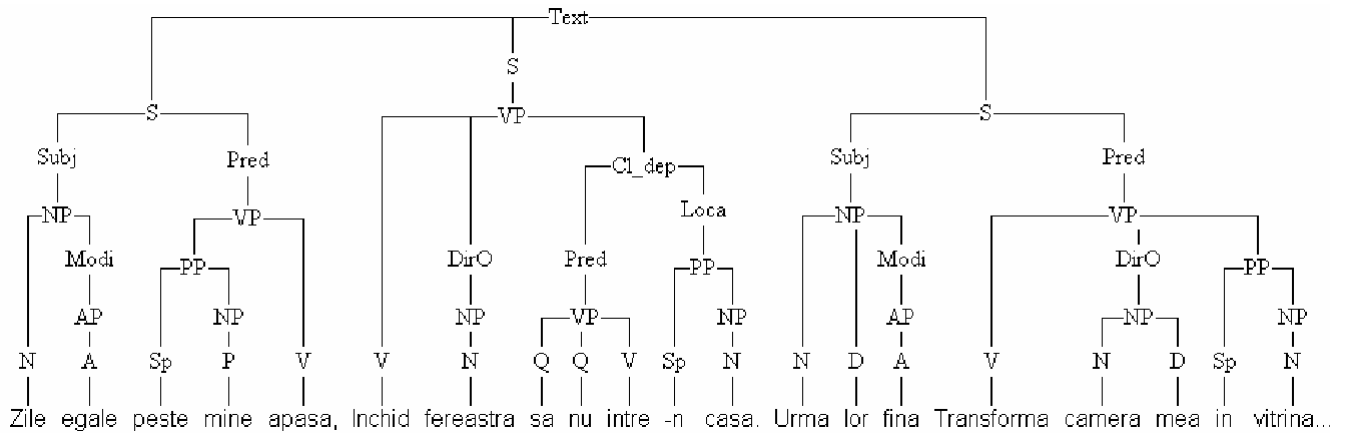


Figura 5.26 a

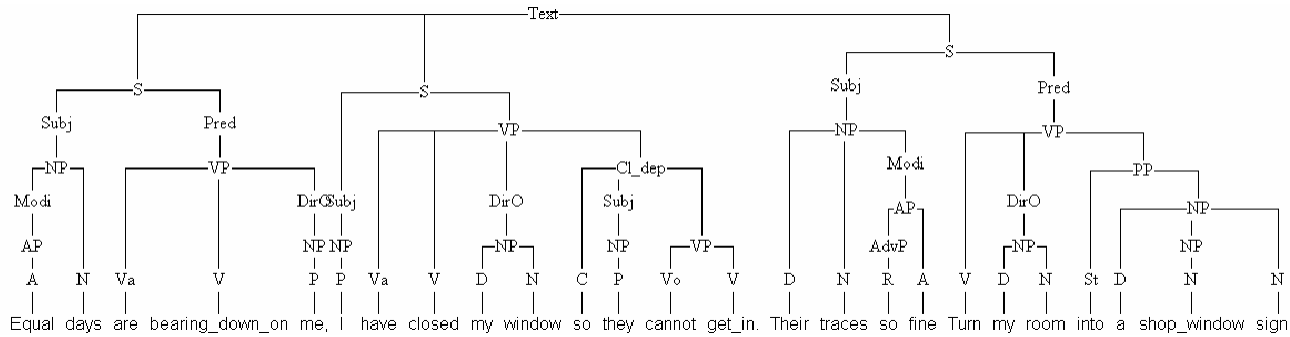


Figura 5.26 b

88 http://www.fantasypieces.net/translation/alexandru_andries/

Textul din Exemplul 20 introduce relația de *codependență sau dependență colaterală*, adică relația dintre trei termeni, A, B, C. Termenul C se realizează numai în prezența termenilor A, B: Eu am întins spre el^A brațul^B ca o ramură^C < *Eu am întins spre el ca o ramură. sau < *Brațul ca o ramură. Același termen C este substituibil cu zero: Eu am întins spre el^A brațul^B ca o ramură^C < Eu am întins spre el^A brațul^B. Termenul C este dependent simultan de A și de B, mai exact, de A, strict sintactic, de B, și semantic, în sensul că C și B prezintă compatibilitate semantică.

Tipar verbal:

[V+ PP [spre [NP_{AC}]]+ NP_{AC} [OD]] [V+ PP [spre [NP_{AC}]]+ NP_{AC} [OD]]_{ro} ☒ [V+ NP_D [OI] + NP_{AC} [OD]] [V+ NP_{AC} [OD] + PP [towards [NP_{AC}]]]_{en}

Exemplul 20:

Eu am întins spre el brațul ca o ramură.

El și-a înclinat spre mine trunchiul
ca un măr.⁸⁹

I offered him my arm like a branch.

He tipped his trunk towards me
like a shoulder.⁹⁰

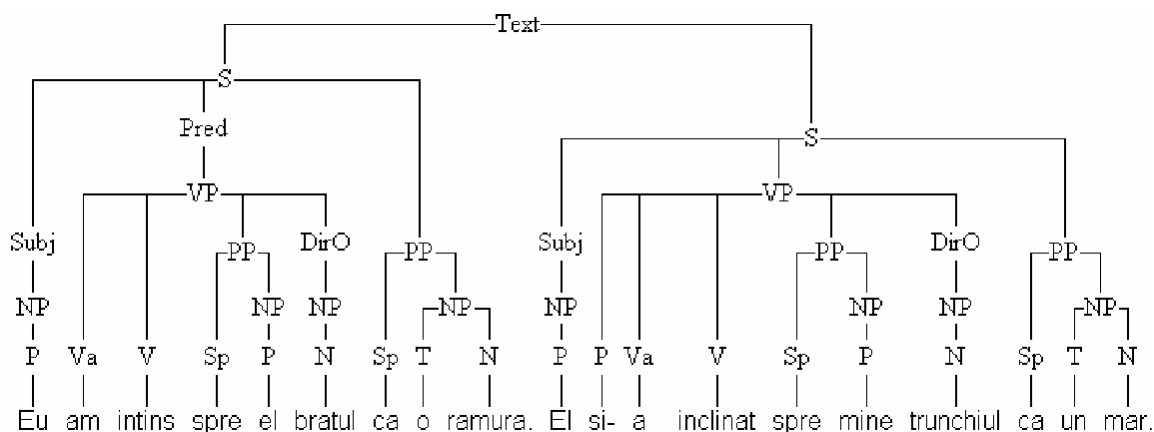


Figura 5.27 a

⁸⁹ <http://poezii.t2i.info/nichita-stanescu/necuvintele/>

⁹⁰ http://www.romanianvoice.com/poezii/poezii_tr/unwords.php

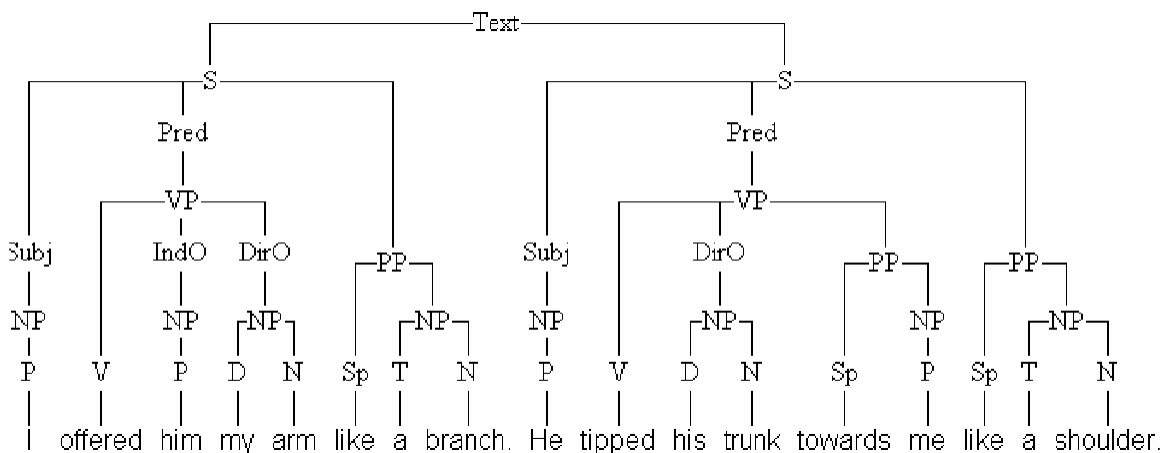


Figura 5.27 b

Relația de echivalență, pe care am definit-o deja, include apoziționarea și dublarea anumitor funcții sintactice. De asemenea, dublarea unor funcții sintactice include: repetiția și corelația acestora, iar corelația presupune reluarea printr-un element anaforic a unui cuvânt sau grup de cuvinte (sintagmă, propoziție, frază sau un segment mixt). Din această perspectivă, pronumele relativ *la care* din Exemplul 21 preia referința de la antecedentul *închinare*, asigură legătura între propoziții, ca atare, este corelativul unui termen antecedent. Fraza "pentru închinarea la care mă supun, / Culeg azur și raze și roze de pe ramuri" este rezultatul legării a două propoziții: "pentru închinarea [...] culeg azur și raze și roze de pe ramuri" și "la închinare mă supun", cu ajutorul pronumelui relativ care are același referent cu substantivul *închinare*.

Tipare verbale:

- [PP [la [NP_{AC}]] + NP_{AC} [OD] + V]_{ro} ☒ [NP_{AC}+ V+ NP_{AC} [OD]]_{en};
- [NP_{AC} [OD]+ V]_{ro} ☒ [NP_{AC} [OD] + PP [with [NP_{AC}]]]_{en}.

Exemplul 21:

pentru închinarea la care mă supun,
 Culeg azur și raze și roze de pe ramuri,
 Stăpânul meu, alesul, cu slavă să-ncunun⁹¹

⁹¹ http://www.fantasypieces.net/translation/vasile_voiculescu/

for the dedication that I bow my head to
 I pick up azure and rays and roses from the branches
 So that my master, that I chose, I can adorn with glory.

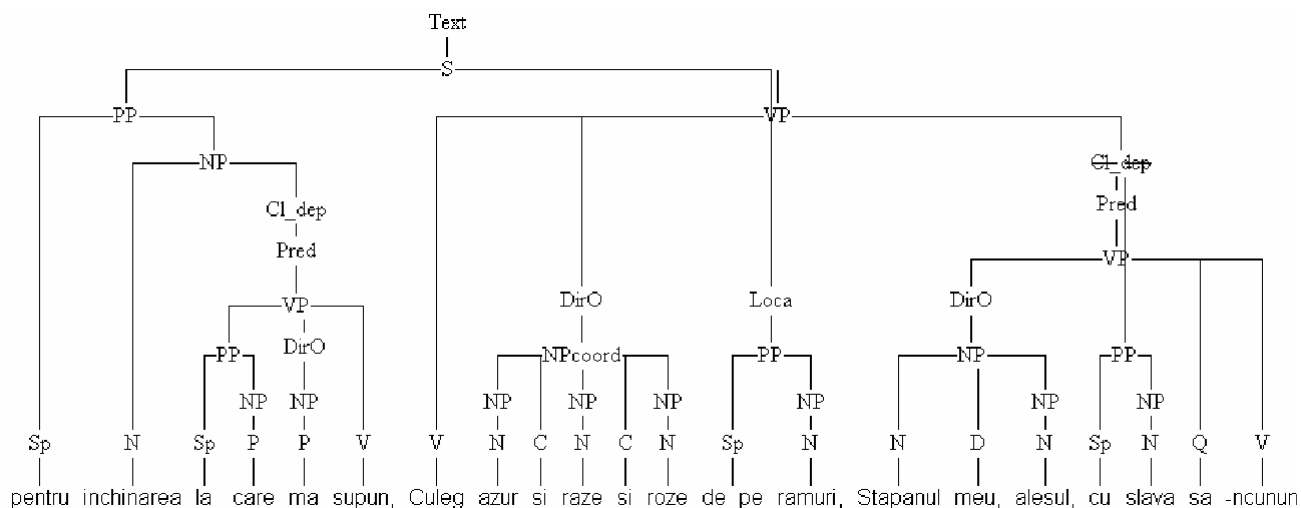


Figura 5.28 a

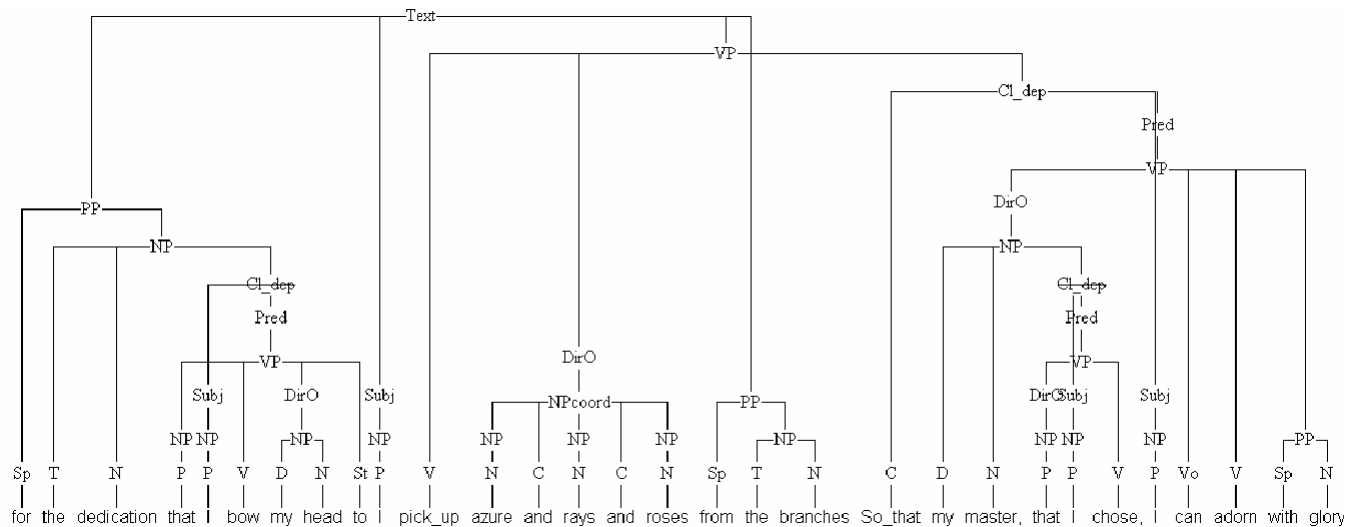


Figura 5.28 b

Structura verbală formată din verb, complement direct, complement prepozițional este ilustrată în Exemplul 22 și în Figurile 5.29 a și b. Complementul direct este anticipat prin cliticul pronominal, în limba română. Complementul prepozițional are aceeași realizare în ambele limbi (prepoziție și grup nominal), iar grupul nominal subcategorizat de prepoziția “cu” include un dativ posesiv exprimat prin forma neaccentuată a pronumelui personal. În limba română acest pronume are funcția de atribut pronominal.

Forma sa este echivalată în limba engleză cu forma adjectivului posesiv, astfel justificându-se alegerea categoriei *Determinant*.

Exemplul 22:

Vă investesc pe voi cu puterea-mi

*I do invest you with my power*⁹²

Tipar verbal:

$[NP_{ObjAgr} + V + PP [pe [NP_{AC}] + PP[Cmpl]^{93}]]_{ro} \boxtimes [V + NP_{AC} + PP[Cmpl]]_{en} =$
 $[PP [pe [NP_{AC}[NP_{ObjAgr} + N]]] + V + PP [Cmpl]]_{ro} \boxtimes [V + NP_{AC} +$
 $PP[Cmpl]]_{en}$

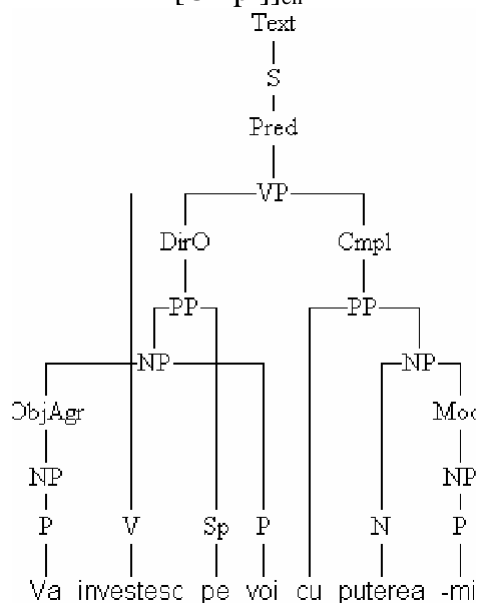


Figura 5.29 a

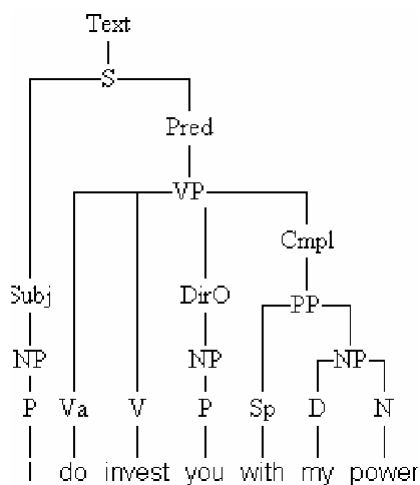


Figura 5.29 b

⁹² W. Shakespeare, op. cit., pp. 52-53

⁹³ Cmpl – eticheta corespunzătoare funcției de Complement

În versurile din textul stănescian **Trist cântec de dragoste**, cele trei grupuri verbale au trei alcătuirii diferite:

- verb și obiect direct exprimat prin substantiv comun inanimat;
- verb, obiect indirect reluat anaforic prin pronume personal, formă neaccentuată de dativ, complement prepozițional introdus prin prepoziția simplă *de*;
- verb și obiect direct exprimat prin clitic pronominal în acuzativ.

Exemplul 23:

Numai iarba știe gustul pământului.
 Numai sângelui meu îi e dor, într-adevăr,
 de inima mea, când o părăsește.⁹⁴

Only the grass knows the taste of the earth.
 Only my blood misses⁹⁵
 my heart when it leaves⁹⁶

Tipar verbal:

[V+ NP_{AC} [OD]] [NP_D [N+ NP_{ObjAgr}]+ V+ PP [de [NP_{AC}]]] [NP_{AC} [N*]+ V]_{ro} ☒ [V+ NP_{AC} [OD]] [NP_{AC} [OD]] [V+ NP_{AC} [OD]]_{en}

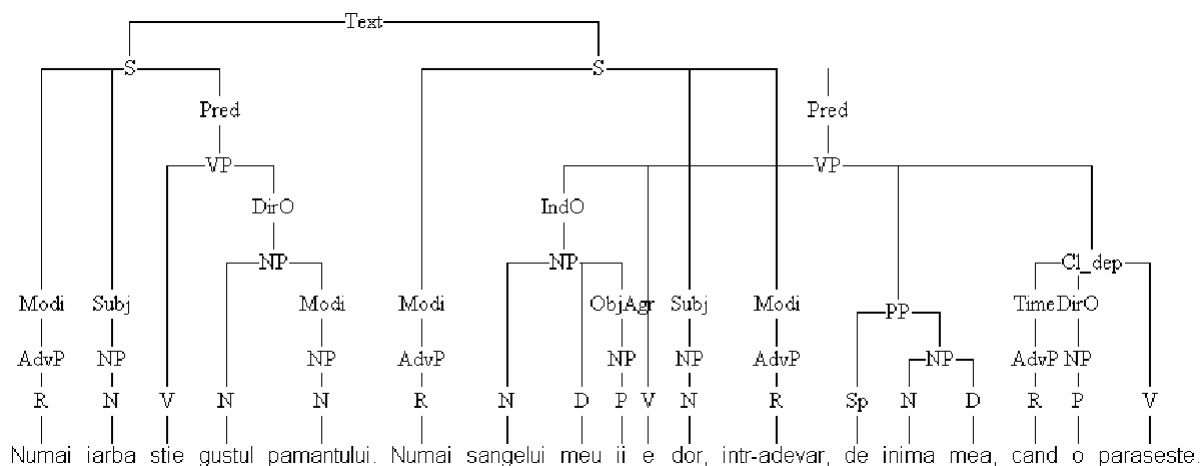


Figura 5.30 a

⁹⁴ http://www.poezie.ro/index.php/poetry/36873/Trist_cantec_de_dragoste

⁹⁵ Verbul *miss* este tranzitiv în acest context (cf. <http://www.thefreedictionary.com/miss>).

⁹⁶ http://www.romanianvoice.com/poezii/poezii_tr/sadlovesong.php

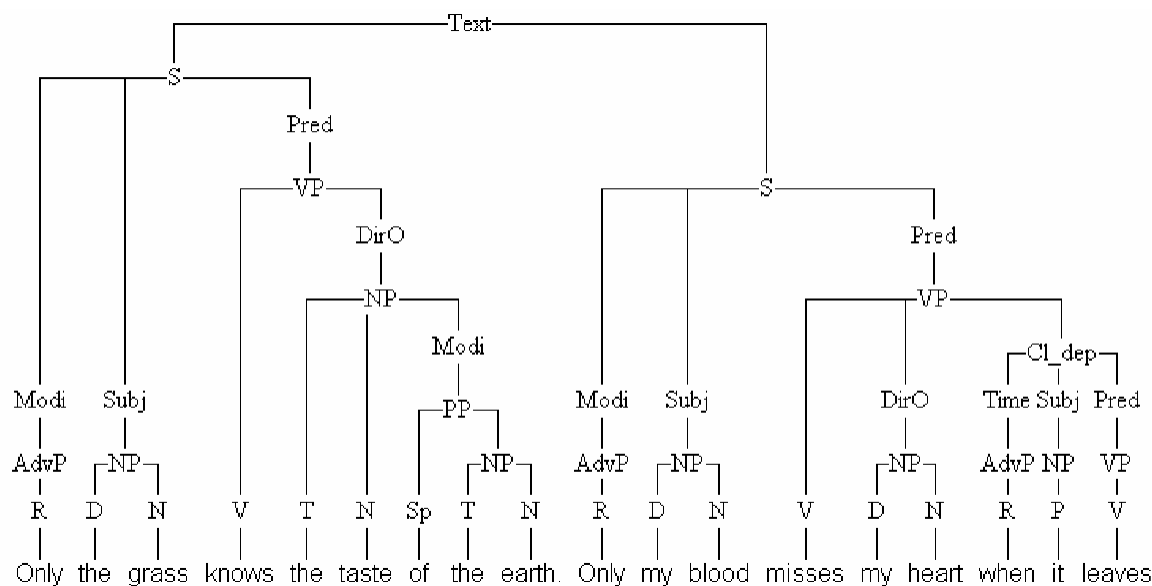


Figura 5.30 b

Accentuarea primului termen al inversiunii *Focul meu a-l stinge nu pot cu toate/Apele mării* conduce la formarea unui grup verbal cu nucleu nepredicativ, în care un obiect direct aflat pe prima poziție este reluat prin clitic pronominal, iar complementul prepozițional este în poziție finală:

Exemplul 24:

Focul meu a-l stinge nu pot cu toate

Apele marii.⁹⁷

Nor can I extinguish my flames with every

Billow of oceans.⁹⁸

Tipar verbal:

[NP_{AC} [N + NP_{ObjAgr}]+ V+ PP [cu [NP_{AC}]]]_{ro} ☒ [V+ NP_{AC} [OD]+ PP [with [NP_{AC}]]]_{en}

⁹⁷ http://www.escoala.ro/referate/romana_oda.html

⁹⁸ http://www.romanianvoice.com/poezii/poezii_tr/ode.php

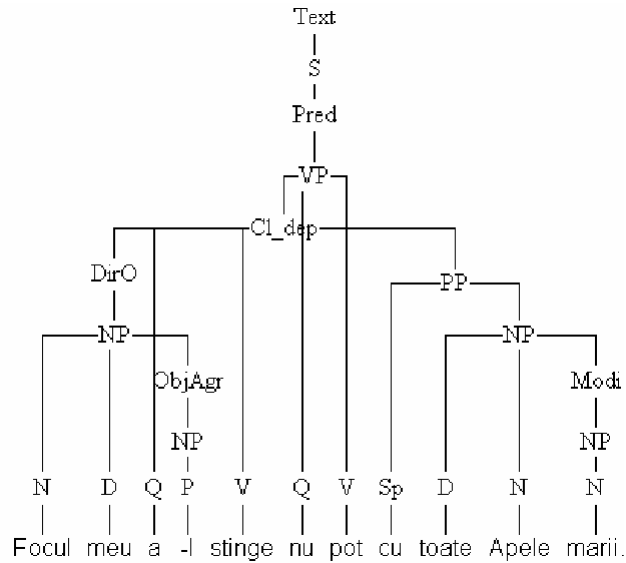


Figura 5. 31 a

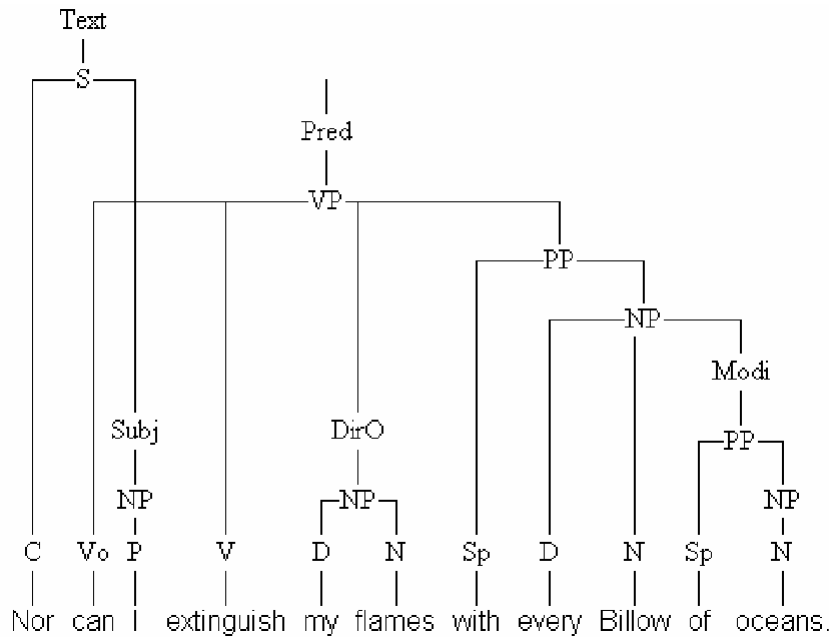


Figura 5. 31 b

Grupul prepozițional *cu toate apele mării* reprezintă o determinare facultativă pentru verbul tranzitiv *stinge* și are rolul sintactic de complement instrumental (*cu apele*). Instrumentul poate fi exprimat sau nu în structura grupului verbal, dar atunci când posibilitatea de selecție este mai mare, actantul facultativ este exprimat.

Distribuția cu nucleu tranzitiv, obiect direct și complement prepozițional poate fi urmărită în Exemplul 25, unde- în afara grupului predicatului, mai există și un grup

prepozițional pe care l-am considerat multiplu și realizat prin coordonare prin juxtapunere⁹⁹. Verbul *a scrie* se definește semantic pe baza mărcilor proprii, fără încorporarea trăsăturilor semantice ale determinanților, iar apariția argumentului din poziție inițială are o explicație de natură stilistică, prin faptul că inversiunea contribuie la realizarea artei poetice *Flori de mucigai*:

Exemplul 25:

Le-am scris cu unghia pe tencuială

Pe un părete de firdă goală,

Pe întuneric, în singurătate

I wrote them with my nail on walls

In the empty alcoves,

In the dark and in solitude¹⁰⁰

Tipar verbal:

[NP_{AC} [N*]+ V+ PP [cu [NP_{AC}]]]_{ro} ☒ [V+ NP_{AC} [N*]+ PP [on [NP_{AC}]]]_{en}

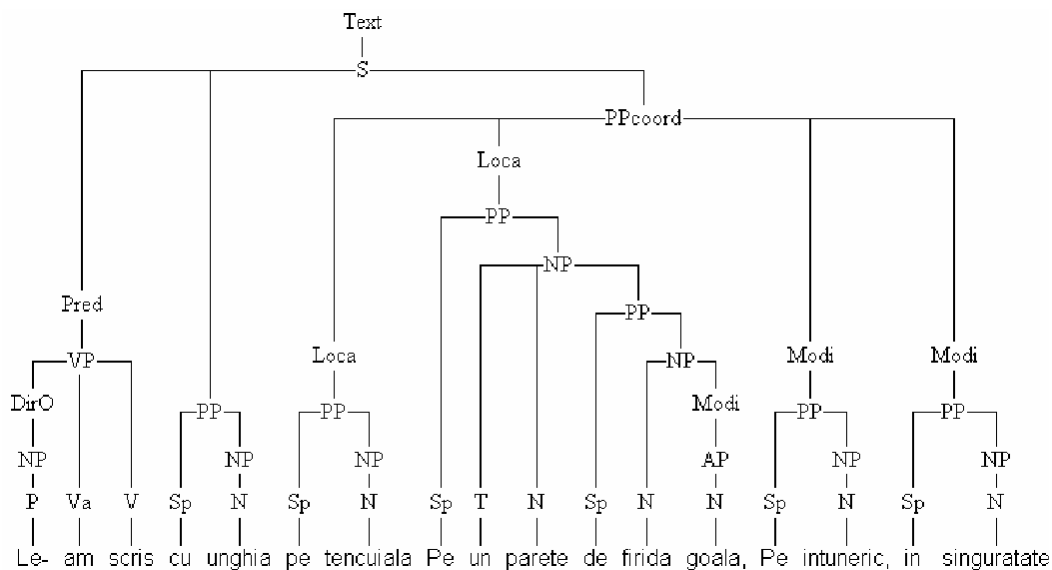


Figura 5. 32 a

⁹⁹ Drept pentru care am definit tipul numit \pt Ppcoord, creat prin relația de coordonare dintre două sau mai multe grupuri de aceeași natură sintactică.

¹⁰⁰ http://www.loritironpandit.com/Loredana_Tiron-Pandit_Web/Written_Translations/Entries/2009/4/11_Mildew_Flowers_by_TUDOR_ARGHEZI.html

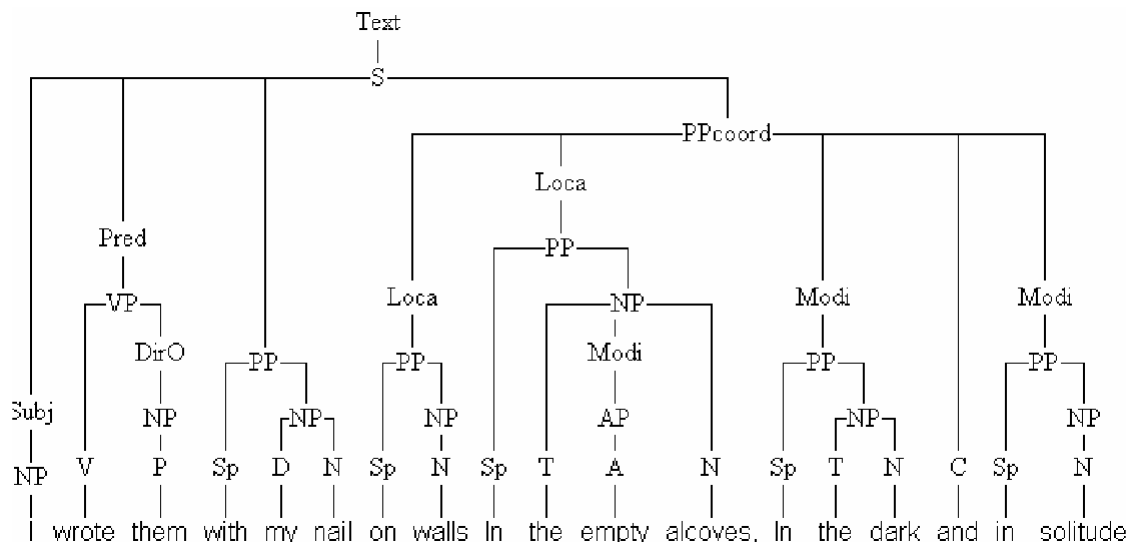


Figura 5.32 b

5.5 Descrierea implementării analizei grupului verbal

Sintaxa reprezintă modulul în care se identifică grupul verbal sau grupurile verbale din fraza de analizat. Pentru acest scop am creat o bază de date distinctă ce conține paradigme de identificare a grupului verbal, *sintaxă.db*.

Din etapa de cercetare a rezultat că numărul maxim de elemente este 7, drept pentru care am stabilit ca baza de date să conțină un număr de 14 coloane, 7 dintre acestea asociate părților de vorbire (de tipul: verb lexical, verb auxiliar, particulă de negație, particulă pentru conjunctiv, nume sau substitut nominal), iar celelalte 7 sunt repartizate părților de propoziție corespondente (care pot fi instanțiate sau nu¹⁰¹). Părțile de vorbire sunt exprimate prin coduri ce reprezintă grupe morfologice. Totalitatea codurilor de pe fiecare linie a bazei de date alcătuiește paradigmele sintagmatice. Acestea, în număr de 116, sunt organizate în anumite combinații poziționale, care au rezultat în urma cercetărilor. Grupele morfologice au următoarea reprezentare în baza de date :

¹⁰¹ De exemplu, substantivul va avea funcția sintactică de complement, dar particula de negație nu va fi asociată, sintactic, niciunei funcții în grupul verbal.

- VB, verb; exemplu: *merg*,
- AX, auxiliar; exemplu: *am, va*
- NO, numeral ordinal; exemplu: *primul, al doilea*
- PA, pronume personal în acuzativ; exemplu: *mă, ne*
- PC, particulă conjunctiv; exemplu: *să*
- PE, prepoziția care introduce complementul direct; prepoziția *pe*
- PF, particula de infinitiv; particula *fi*
- PG, pronume negativ; exemplu: *nimic, niciunul*
- PH, pronume nehotărât; exemplu: *fiecare, altul*
- PN, particula pentru negație *nu*
- PS, pronume posesiv; exemplu: *a mea, al meu*
- SA, substantiv comun [+animat]; exemplu: *vecinul*
- SC, substantiv [-animat]; exemplu: *rochie*

Principiul de funcționare al acestui modul se bazează pe compararea codurilor morfologice corespondente cuvintelor existente în frază cu cele din baza de date. În acest scop se ține cont de pozițiile acestora și de faptul că nu este obligatoriu ca toate codurile unei paradigme să se regasească în frază pentru ca un grup verbal să fie validat. Se va considera valid rezultatul cu cel mai mare număr de elemente identificate. Pașii care fac aplicația funcțională și eficientă sunt:

a. definim un vector de șiruri de caractere, *e*, ce poate conține între 2 și 7 elemente (*e: array [2..7] of strâng;*) și care va primi ca valori coduri morfologice din baza de date începând cu cea de a doua coloană;

b. creăm două liste de tip *tstringlist*, în care se vor adăuga, în urma unei verificări în lista cu coduri sintactice corespunzătoare cuvintelor din fraza, respectiv traducerilor acestora, doar elementele cu cod sintactic efectiv și traducerile corespondente;

c. se inițializează baza de date;

d. verificăm existența primului cod sintactic din listă în prima coloană a bazei de date. Dacă există, se filtrează baza de date după acest cod limitându-se semnificativ numărul de paradigme rămase pentru verificări; în caz contrar, se caută în următoarea coloană;

e. imediat după filtrare, vectorului 'e' îi sunt atribuite restul de elemente ale paradigmei curente;

f. se compară, pe rând, următorul element din lista ce conține codurile sintactice, denumită 'ls', cu fiecare din elementele atribuite vectorului 'e' până când apare o egalitate, moment în care se trece la compararea următorului element din lista 'ls' cu restul de elemente ce aparțin vectorului 'e'. Acest fenomen se repetă până la epuizarea elementelor vectorului 'e';

g. întregul proces de la punctul (f) se reia pentru fiecare din paradigmele rămase după filtrare, cu excepția cazului în care la un moment dat sunt identificate într-o paradigmă toate codurile sintactice din lista 'ls'. În acest moment grupul verbal este validat, iar scopul acestui modul este finalizat, trecându-se la afișarea grupului verbal respectiv;

h. pentru fiecare egalitate între codurile sintactice ale cuvintelor și cele existente în paradigma curentă, am creat o variabilă denumită 'gvt'(grup verbal temporar) căreia îi sunt atribuite traducerile corespunzătoare codurilor sintactice. La terminarea verificării primei paradigme, valoarea variabilei 'gvt' este transferată variabilei ce va reprezenta la final grupul verbal, și anume 'gvf'(grup verbal final). După fiecare din paradigmele următoare cele două variabile sunt comparate, variabila 'gvf' păstrând valoarea cea mai mare.

i. la finalul verificării paradigmelor, se validează grupul verbal, valoarea lui fiind dată de valoarea variabilei 'gvf' și se afișează.

Capitolul 5

Concluzii

În introducere am propus ideea interpretării limbajului ca interacțiune între lexicon și gramatică, am definit grupul verbal și am stabilit structura sa sintactică din perspectiva trăsăturii de tranzivitate. După care, în celelalte secțiuni am descris din perspectivă teoretică și practică modulele aplicației de traducere automată AGV1.0, și anume: deflexionarea nominală și verbală pentru limba sursă, relația cu modulul de preprocesare a textelor TTL¹⁰², flexionarea nominală și verbală pentru limba țintă, compunerea grupului verbal.

Contribuțiile personale se regăsesc în cele două metode de deflexionare pe care le-am propus, ambele bazate pe morfologia paradigmatică, pe restricții de unificare de șir și pe operații de modificare de șir (*adun șir*, *scad șir*, *alternanță șiruri*). De asemenea, contribuții personale înseamnă și organizarea, respectiv conducerea studiului de analiză sintactică, prin realizarea arborilor sintactici și a tiparelor verbale asociate grupului verbal cu nucleu tranzitiv.

La nivel practic, contribuțiile personale implică proiectarea și crearea bazelor de date (pentru deflexionare / flexionare; pentru sintaxă) și algoritmi de identificare a formelor lemă, de creare a formelor flexionate, respectiv de compunere a grupului verbal și de analiză a constituenților acestuia.

Lucrarea **Naturațea limbajului artificial - studiu despre traducerea automată a GV** reprezintă o experiență care are nevoie de un punct terminus pentru a trece la un alt nivel, din perspectiva căruia să fie reorganizate conceptele și modalitățile de lucru. O direcție viitoare de lucru poate fi aceea de a demonstra că limbajul natural este un sistem complex adaptiv [25, 27, 22, 23], păstrând modul de realizare a grupului verbal, dar receptându-l dintr-o altă dimensiune.

A afirma că un grup verbal are comportamentul unui sistem adaptiv complex (CAS) implică următoarele aspecte fundamentale:

¹⁰² În engleză, “Tokenizing, Tagging and Lemmatizing free running texts”.

- șabloane de activitate: cercetătorul poate descrie structura lexical conceptuală pentru fiecare verb considerat intrare lexicală într-o bază de date multilingvă, considerând structura lexical conceptuală drept un șablon de reprezentare sintactico-semantică a grupului verbal;
- auto-organizare: structura primește o interpretare holistică, incluzând un tip, o primitivă și un câmp¹⁰³;
- comportament colectiv: nucleul verbal este extins, astfel încât rădăcina unei structuri lexical conceptuale acceptă toate verbele care aparțin aceleiași clase semantice.

În concluzie, ceea ce îmi propun în continuarea acestei lucrări este o argumentare a interpretării grupului verbal în termenii unui sistem adaptiv complex, plecând de la o definiție scurtă, dar foarte clară dată de Plsek, Lindberg și Zimmerman¹⁰⁴: „Un Sistem Adaptiv Complex este un sistem compus din agenți individuali, care au libertatea de a acționa în moduri care nu sunt total predictibile și ale căror acțiuni sunt interconectate astfel încât acțiunile unui agent schimbă contextul pentru alți agenți.”

¹⁰³ În lucrarea lui A.N. Fazil și B.J. Dorr [33], o structură lexical conceptuală (LCS) este un graf orientat cu rădăcină, astfel încât fiecare rădăcină este asociată cu o informație particulară, incluzând *tipul*, *primitiva* și *câmpul*. Numele de tipuri sunt *Event*, *Path*, *Manner*, *Property*, *Thing*; câmpurile se referă la valorile *Locational*, *Possessional*, *Identificational*. Primitiva unui nod LCS este, alcătuită dintr-o primitivă structurală (de exemplu, *go*, *cause*, *act*) și constante (de exemplu, *reduce* + *ed*, *slash* + *ingly* pentru limba engleză; *câine* + *ește*, *face* + *ut* pentru limba română).

¹⁰⁴ Plsek, P., Lindberg, C., Zimmerman, B. (1997) **Some Emerging Principles for Managing Complex Adaptive Systems**. Working Paper version Nov. 25, p. 2

Bibliografie

- [1] Academia Română, Institutul de Lingvistică “Iorgu Iordan- Al. Rosetti” (2005) **Gramatica Limbii Române**, vol. I și II, Editura Academiei Române, București.
- [2] Anick, P., Artemieff, S. (1992) **A high-level morphological description language exploiting inflectional paradigms**, *Proceedings of the 14th conference on Computational linguistics* - Volume 1, Pages: 67 – 73, Morristown.
- [3] Avram, M. (1997) **Gramatica pentru toți**, ediția a II-a, Editura Umanitas, București.
- [4] Avram, M. (2001) **Gramatica pentru toți**, ediția a III-a, Editura Humanitas, București.
- [5] Barbu, A. M., Ionescu, E. (1996) **Teorii gramaticale contemporane: Gramatica Centrilor de Sintagmă**, *Limba Română*, 1, pp: 31-55.
- [6] Berent, I., Pinker, S., & Shimron, J. (1999) **Default nominal inflection in Hebrew: Evidence for mental variables**. *Cognition*, 72, pp. 1-44.
- [7] Black, Ch. A. (1996) **Introduction to Government and Binding theory: X-Bar theory applied to languages with other word orders**, *Notes on Linguistics* 75, pp: 5-16.
- [8] Bulgăr, Gh. (1995) **Limba română. Fonetică, lexic, morfologie, sintaxă, stilistică**, Editura Vox, București.
- [9] Calder, J. (1989) **Paradigmatic morphology**. *Proceedings of the fourth conference on European chapter of the Association for Computational Linguistics*, Manchester.
- [10] Carl, M., Way, A. (eds.), (2003) **Recent Advances in Example-Based Machine Translation**, Kluwer Academic Publishers, Printed in Netherlands
- [11] Carl, P., Sag, I. (1987) **An Information-based Syntax and Semantics**, volume 13. CSLI Lecture Notes, Stanford University.
- [12] Chiorean, L. (2009) **Adnotări la gramatica prepoziției**, în revista *Limba Română*, nr. 11-12, anul XIX, Chișinău.
- [13] Codrin, Ș. (1994) **Dincolo de tăcere**, Editura Haiku, București.
- [14] Collins, M., Brooks, J. (1995) **Prepositional phrase attachment through a backed-off model**. In *Proceedings of the Third Workshop on Very Large Corpora*, Cambridge, pp. 27–38.

- [15] Constantinescu-Dobridor Gh. (2001) **Gramatica limbii române**, Editura Didactică și Pedagogică, București, p. 134.
- [16] Corbe, M., Tabory, R., **Introduction to an Automatic English Syntax (by fragmentation)**.
- [17] Coseriu, E., **Pour et contre l'analyse sémique**, Tokio, p. 137.
- [18] Curteanu, N., Trandabăț, D., Moruz, M. (2006) **Structura grupului verbal, reprezentarea logică a predicatului și predicția lexicală în limba română**, in C. Forăscu, D. Tufiș, D. Cristea (eds.) *Proceedings of the Workshop Linguistic Resources and Tools for Processing Romanian Language*, Iași.
- [19] Crystal, D. (1980) **A first dictionary of linguistics and phonetics**, Boulder, CO: Westview.
- [20] Dincă (Huțuliac), N.L. (2006) **Frames for a PC-PATR implementation of the Romanian transitive verbs**, in C. Forăscu, D. Tufiș, D. Cristea (eds.) *Proceedings of the Workshop Linguistic Resources and Tools for Processing Romanian Language*. University "Al.I. Cuza" Publishing House, Iași.
- [21] Dincă (Huțuliac), N.L. (2007) **Thematic roles from subcategorization schema of transitive verbs**, *Proceedings of the International Conference "Language and Literature- European Landmarks of Identity"*. University "C^{tin}. Brancusi" Publishing House, Pitești.
- [22] Dincă, N.L. (2007) **Learning Definiend Concept and Possible Worlds of Verbal Phrase**, in „Studii și cercetări filologice. Seria Limbi străine aplicate”, nr. 6, 2007, University "C^{tin}. Brancusi" Publishing House, Pitești.
- [23] Dincă, N.L. (2007) **Morphological analysis proposal based on nominal paradigms**, in I. Pistol, D. Tufiș, D. Cristea (eds.) *Proceedings of the Workshop Linguistic Resources and Tools for Processing Romanian Language*. University "Al.I. Cuza" Publishing House, Iași.
- [24] Dincă, N.L. (2007) **Consideratii teoretice asupra aplicabilitatii unei baze de date cu exemple de traducere**, in D. M. Trandabat, D. Cristea, D. Tufiș (eds.) *Proceedings of the Workshop Linguistic Resources and Tools for Processing Romanian Language*. University "Al.I. Cuza" Publishing House, Iași.

- [25] Dincă, N.L. (2008) **Modeling a Quite Different Machine Translation from Lexical Conceptual Structure** in “Polibits”, Journal of Research and Development in Computer Science and Engineering, Number 38, Mexico.
- [26] Dincă, N.L. (2008) **Rule-Based Morphological Disambiguation**, in „Studii de gramatică contrastivă”, nr. 8, University “C^{tin}. Brancusi” Publishing House, Pitești.
- [27] Dincă, N.L. (2008) **Word Structure**, in „Studii de gramatică contrastivă”, nr. 9, University “C^{tin}. Brancusi” Publishing House, Pitești.
- [28] Dincă, N.L. (2008) **The English verb, behind regular – irregular distinction** in „Studii și cercetări filologice. Seria Limbi străine aplicate”, nr. 7, University “C^{tin}. Brancusi” Publishing House, Pitești.
- [29] Dincă, N.L. (2008) **An Introduction to Example-Based Machine Translation**, Proceedings of the International Conference “Language and Literature- European Landmarks of Identity”. University “C^{tin}. Brancusi” Publishing House, Pitești.
- [30] Dincă, N.L. (2010) **Example Database Design, Conceived as a Synergetical Research Direction**, in “Comunicare profesională și traductologie”. Editori: Conf. dr. Rodica SUPERCEANU, asist. drd. Daniel Dejica, 2010, Volum 3, ISBN 973-625-175-6. Editura Politehnica, Timișoara.
- [31] Dorr, B.J., Jones, D. (1996) **Role of word sense disambiguation in lexical acquisition: predicting semantics from syntactic cues**, *Proceedings of the 16th conference on Computational linguistics*, Copenhagen, pp: 322 – 327.
- [32] Dorr, B. J., Green, R., Resnik, Ph. (2004) **Inducing Frame Semantic Verb Classes from WordNet and LDOCE**, in *Proceedings of the Association for Computational Linguistics*, Barcelona, pp: 96-102.
- [33] Dorr, B. J., Fazil, A. N. (2002) **Generating A Parsing Lexicon from an LCS-Based Lexicon**, in *Proceedings of the LREC-2002 Workshop on Linguistic Knowledge Acquisition and Representation*, Las Palmas, pp: 43—52.
- [34] Dorr, B. J. (Mari Olsen, Nizar Habash, and Scott Thomas), **LCS Verb Database**, Online Software Database of Lexical Conceptual Structures and Documentation, UMCP, 2001.

- [35] Dorr, B. J., Fazil, A. N. (2002) **Generating A Parsing Lexicon from an LCS-Based Lexicon**, in *Proceedings of the LREC-2002 Workshop on Linguistic Knowledge Acquisition and Representation*, Spain, pp. 43-52.
- [36] Dorr, B. J., Nizar, H., Traum, D. (2003) **Hybrid Natural Language Generation from Lexical Conceptual Structures**, *Machine Translation*, 18:2, pp. 81—128.
- [37] Ducháček, (1967) **Précis de sémantique française**, Brno, p. 20.
- [38] Fisticanu, T. (2007) **Mutații semantice în cadrul îmbinărilor sinestezice din limbajul poetic**, *Limba română*, nr. 7-9 (disponibil la adresa: <http://www.limbaromana.md/index.php?go=articole&printversion=1&n=705>).
- [39] Forăscu, N., Popescu, M. (2001) **Dificultăți gramaticale ale limbii române**, Editura Universității, București.
- [40] Gălățeanu, G., Comișel, E., **Gramatica limbii engleze**, Editura Didactică și Pedagogică, București, 1982.
- [41] Grimshaw, J. (1990) **Argument Structure**. MIT Press.
- [42] Guthrie, J.A., Guthrie, L., Wilks, Y., Aidinejad, H., **Subject-dependent co-occurrence and word sense disambiguation**, *Proceedings of the 29th annual meeting on Association for Computational Linguistics*, Berkeley, pp: 146 – 152.
- [43] Guțu Romalo, V. (1963) **În problema clasificării verbelor. Încercare de clasificare sintagmatică**, SCL, nr.1, pp. 29-41.
- [44] Guțu Romalo, V. (1973) **Gramatica limbii române**, EDP, București.
- [45] Hindle, D., Rooth, M. (1993) **Structural ambiguity and lexical relations**. *Computational Linguistics*, 19(1), pp:103–120.
- [46] Hjelmslev, L. (1967) **Preliminarii la o teorie a limbii**, trad. rom. D. Copceag, București.
- [47] Iorgu, I., Guțu-Romalo, V, Niculescu, Al. (1967) **Structura morfologică a limbii române**, Editura Științifică, București, p. 181.
- [48] Irimia, D. (1997) **Gramatica limbii române**, Editura Polirom, Iași.
- [49] Karlsson, F. (1990) **Constraint Grammar as a Framework for Parsing Unrestricted Text**. *Proceedings of the 13th International Conference of Computational Linguistics*, Vol. 3, pp. 168-173, Helsinki.

- [50] Karlsson, F. (2005) **Nature and methodology of grammar writing**. *SKY Journal of Linguistics* 18, 2005, pp. 341-356.
- [51] Karov, Y., Edelman, S. (1998) **Similarity-based word sense disambiguation**, *Computational Linguistics*, v.24 n.1, pp: 41 – 59.
- [52] Karttunen, L., Kaplan, R. M., Zaenen, A. (1992) **Two-level morphology with composition**, *Proceedings of the 14th conference on Computational linguistics - Volume 1*, pp: 141 – 148, Morristown.
- [53] Koktová, E. (1985). **Towards a new type of morphemic analysis**. *Proceedings of the second conference on European chapter of the Association for Computational Linguistics*, Geneva.
- [54] Koskenniemi, K. (1984) **A general computational model for word-form recognition and production**. *Proceedings of the 22nd annual meeting on Association for Computational Linguistics*, Stanford.
- [55] Lazăr, L. (2007) **Limbaj și semnificație: distincția „semantic vs. real” în lingvistica integrală**, în revista *Limba Română*, nr. 4-6, anul XVII, Chișinău.
- [56] Leech, G., Wilson, A. (1994) **EAGLES morphosyntactic annotation**. *EAGLES report EAGSCSG/IR-T3.1*. Pisa: Istituto di Linguistica Computazionale.
- [57] Levin, B. (1993) **English Verb Classes and Alternations - A Preliminary Investigation**, 1993, The University of Chicago Press.
- [58] Linguistic Tree Constructor: <http://ltc.sourceforge.net/>
- [59] KIMMO: <http://www.sil.org/pckimmo/pc-kimmo.html>
- [60] Marcus, S., Nicolau, E., Stati, S. (1966) **Introducere în lingvistica matematică**, vol. I, pp. 26-41.
- [61] Manning, Ch., Schütze, H. (1999) **Foundations of Statistical Natural Language Processing**, MIT Press. Cambridge.
- [62] McCarthy, D., Korhonen, A. (1998) **Detecting verbal participation in diathesis alternations**, *Proceedings of the 36th annual meeting on Association for Computational Linguistics*, Montreal, pp: 1493 – 1495.
- [63] Merlo, P., Ferrer, E. E. (2006) **The Notion of Argument in Prepositional Phrase Attachment**, *Computational Linguistics*, Volume 32, Issue 3, MIT Press, pp: 341 – 378.

- [64] Merlo, P., Crocker, M., Berthouzoz, C. (1997) **Attaching multiple prepositional phrases: Generalized backed-off estimation**. In *Proceedings of the Second Conference on Empirical Methods in Natural Language Processing*, Providence, pp: 145–154.
- [65] Merlo, P., Leybold, L. (2001) **Automatic distinction of arguments and modifiers: The case of prepositional phrases**. In *Proceedings of the Fifth Computational Natural Language Learning Workshop (CoNLL-2001)*, Toulouse, pp: 121–128.
- [66] Mierlă, A. (2003) **Considerații privind criteriile de clasificare sintactico-semantică a verbelor (cu aplicație asupra verbelor de deplasare)**. *Analele Științifice ale Universității Ovidius, Constanța, Seria Filologie*, nr. 14, pp. 201-212.
- [67] Nerbonne, J. (1993) **Feature-based allomorphy**. *Proceedings of the 31st annual meeting on Association for Computational Linguistics*, Columbus.
- [68] Pană Dindelegan, G. (1974) **Sintaxa transformațională a grupului verbal în limba română** București, Editura Academiei, pp. 19-20.
- [69] Pană Dindelegan, G. (1973) **Grup prepozițional facultativ vs. grup prepozițional obligatoriu**, SCL nr. 5.
- [70] Pană Dindelegan, G. (1976) **Reflecții asupra modalității contextuale de analiză a sensului (cu referire specială la verb)**, LR, nr. 2.
- [71] Pană Dindelegan, G. (1978) **Determinant obligatoriu vs. determinant facultativ; o paralelă sintactică între verb și substantiv**, SCL, nr. 5.
- [72] Newmark, P. (1988) **Approaches to Translation**. London: Prentice Hall.
- [73] Pinker, S., Prasada, S. (1993) **Generalizations of regular and irregular morphology**. *Language and Cognitive Processes*, 8 , pp: 1-56.
- [74] Pinker, S. (1991) **Rules of Language**, *Science*, 253, pp: 530-535.
- [75] Pinker, S., Ullman, M. (2002) **The past and future of the past tense**. *Trends in Cognitive Science*, 6, pp.: 456-463.
- [76] Pinker, S. (2000) **The Irregular Verbs**. Landfall.
- [77] Pinker, S. (2003) **Language as an adaptation to the cognitive niche**. In M. Christiansen & S. Kirby (Eds.), *Language evolution: States of the Art*. New York: Oxford University Press.

- [78] Pirrelli, V., Federici, S. (1994) **"Derivational" paradigms in morphonology**. *Proceedings of the 15th conference on Computational linguistics*, Kyoto.
- [79] Popescu, Șt. (1995) **Gramatica practică a limbii române**, ediția a IV-a, Editura Didactică și Pedagogică, București, p. 240.
- [80] Ritchie, G. (1992) **Languages generated by two-level morphological rules**, *Computational Linguistics*, v.18 n.1, p.41-59.
- [81] Saint-Dizier, P. (ed.) (1999) **Predicative forms in natural language and in lexical knowledge bases**. Dordrecht: Kluwer Academic Publishers.
- [82] Sarkar, S., Tripasai, W. (2002) **Learning verb argument structure from minimally annotated corpora**, *Proceedings of the 19th international conference on Computational linguistics*, pp.1-7, Taipei.
- [83] Satoshi, S., Nagao, M. (1990) **Toward Memory- Based Translation**, in *Proceedings of the 13th conference on Computational linguistics*, Finland: Helsinki, 1990, pp. 247-252.
- [84] Schütze, C. T. (1995) **PP Attachment and Argumenthood**. *MIT Working Papers in Linguistics*, 26, pp:95–151.
- [85] Ferdinand de Saussure (1998) **Curs de lingvistică generală**, Editura Polirom, Iași, p. 99.
- [86] Shakespeare, W. (2000) **Regele Lear**, Ediție bilingvă, Institutul European.
- [87] Schulte, S. (2000) **Clustering verbs semantically according to their alternation behaviour**, *Proceedings of the 18th conference on Computational linguistics*, Saarbrücken, pp: 747 – 753.
- [88] Stati, S. (1972) **Elemente de analiză sintactică**, Editura Didactică și Pedagogică, București.
- [89] Șerban, V. (1974) **Teoria și topica propoziției în româna contemporană**, Editura Didactică și Pedagogică, București, pp. 45-47.
- [90] Tanaka, H. (1994) **Verbal case frame acquisition from a bilingual corpus: gradual knowledge acquisition**, in *Proceedings of the 13th conference on Computational linguistics*, Kyoto, pp: 727-731.
- [91] Trăușan Matu, St.(2001). www.racai.ro/~trausan/morfo2.pdf

- [92] Tufiș, D., Andersen, P. (eds.), (1997). **Recent Advances in Romanian Language Technology**. Editura Academiei Române, București, 1997.
- [93] Tufiș, D., Popescu, O. (1991) **A unified management and processing of word-forms, idioms and analytical compounds**. *Proceedings of the fifth conference on European chapter of the Association for Computational Linguistics*, Berlin.
- [94] Tufiș, D. (1989) **It Would Be Much Easier If WENT Were GOED**. In Harry Somers and Mary McGee Wood (eds.), *Proceedings of the 4th European Conference of the Association for Computational Linguistics* (EACL 1989), pp. 145-152, Manchester.
- [95] Tufiș, D., Diaconu, L., Barbu, A.M., Diaconu, C. (1996) **Morfologia limbii române, o resursă lingvistică reversibilă și reutilizabilă**. In Dan Tufiș (ed.), **Limbaj și Tehnologie**, pp. 59-66. Editura Academiei Române, București, 1996.
- [96] Tufiș, D., Barbu, A. M., Pătrașcu, V., Rotariu, G., Popescu, C. (1997) **Corpora and Corpus-Based Morpho-Lexical Processing**. In Dan Tufiș and Poul Andersen (eds.), *Recent Advances in Romanian Language Technology*, Editura Academiei Române, pp. 35-56.
- [97] Tufiș, D., et al. (2008) **RACAI's Linguistics Web Services**, in *Proceedings of the 6th Language Resources and Evaluation Conference - LREC 2008*, Marrakech, Morocco, May 2008.
- [98] Tufiș, D. (1999) **Tiered Tagging and Combined Language Models Classifiers**. in Václav Matousek, Pavel Mautner, Jana Ocelíková, and Petr Sojka, editors, **Text, Speech and Dialogue** (TSD 1999), Lecture Notes in Artificial Intelligence 1692, pp. 28-33. Springer Berlin / Heidelberg, January 1999.
- [99] Tufiș, D. et al. (2010) **Servicii web interoperabile și multilinguale**, in Adrian Iftene, Horia N. Teodorescu, Dan Cristea, Dan Tufiș (eds.) (2010), *Resurse lingvistice și instrumente pentru prelucrarea limbii române*, Ed. Universității Al. I. Cuza, pp. 167-176.
- [100] Ungureanu, V. (2005) **Verbele tranzitive în limba română**, CNAA (teză de doctor în filologie), Chișinău.
- [101] Vasiliu, E. (1963) **Gramaticile generative**, Limba română, nr. 3, pp. 219-231

- [102] Vasiliu, E., Golopenția Eretescu, S. (1969) **Sintaxa transformațională a limbii române**, pp. 15-37
- [103] Vasiliu, F. (1993) **Tolba cu licurici**, Editura Haiku, București.
- [104] Visdic: <http://nlp.fi.muni.cz/projekty/visdic/>
- [105] XEROX: <http://www.xrce.xerox.com/research/mltt/>

ANEXE

TABEL REGULI DEFLEXIONARE NOMINALĂ - LIMBA ROMÂNĂ

Restricție de verificat	Identificatori				Gen.- Dat.	N-Ac pl. art hotarat		N- Ac,pl.near	Gen.Dativ,sg,art.hotarat			Nom-Acuzativ, sg., art hot			Nom- Ac.sg.,near.	
					term	scad sir	adun sir	scad sir	alternante	scad sir	adun sir	scad sir	altern	adun sir	scad sir	adun sir
terminatie																
âni, uri	lor	i	lui	le	lor	lor	i	i	i<e		lui	lui		le	le	
ieți	lor	i	ului	ul	lor	lor	i	i	ț<t, ie<ia	i	ului	ui			ul	
și, ști	lor	i	ului	ul	lor	lor	i	i	ș<s	i	ului	ui			ul	
chi, poi	lor	i	ului	ul	lor	lor	i	i			ului	ui			ul	
ați	lor	i	ălui	l	lor	lor	i	i	ț<t	i	ălui	ui			l	
ii,cei,cai	lor	i	lului	ul	lor	lor	i	i		i	lului	lui			ul	
nți	lor	i	lui	le	lor	lor	i	i	i<e,ț<t		lui	lui		le	le	
poi	lor	i	ului	ul	lor	lor	i	i			ului	ui			ul	
ri	lor	i	ului	ul	lor	lor	i	i	er<ăr	i	ului	ui			ul	
ței	lor	i	lului	ul	lor	lor	i	i		i	lului	ui			ul	
zi	lor	i	ului	ul	lor	lor	i	i	z<d	i	ului	ui			ul	
oi, ei	lor	i	ului	l	lor	lor	i	i		i	ului	ui			l	
ieni	lor	i	ului	ul	lor	lor	i	i	ie<ia	i	ului	ui			ul	
eni	lor	i	ului	ul	lor	lor	i	i	e<ea	i	ului	ui			ul	
eri	lor	i	ului	ul	lor	lor	i	i	i<ă, e<ă	i	ului	ui			ul	
eți	lor	i	elui	le	lor	lor	i	i	ț<t	i	elui	lui			le	
gii	lor	i	ului	ul	lor	lor	i	i		i	ului	ui			l	
ori	lor	i	ului	le	lor	lor	i	i	o<oa	i	elui	ui		e	le	
ți,enți	lor	i	ului	ul	lor	lor	i	i	ț<t	i	ului	ui			ul	
neregulat: oameni																
(et,eț,ep,erz,etr, est)_e	lor	le	i	a	lor	lor	le	le			i	ei	e<a	a	a	ă
le	lor	le	i	ua	lor	lor	le	le			i	lei		ua	ua	
âni	lor	le	i	a	lor	lor	le	le			i	ii	âi<â	a	a	ă
ărți,ăți	lor	le	i	a	lor	lor	le	le			i	ii	ă<a,ț<t	ea	a	
oi	lor	le	i	a	lor	lor	le	le			i	i	o<oa	a	a	e
ci,uri,imi,ni	lor	le	i	a	lor	lor	le	le			i	ii		a	a	ă

ști	lor	le	ei	a	lor	lor	le	le	șt<st	i	ei	ei		ea	a	
ăni, ărci, ălci, ăși, ărbi, cări	lor	le	i	a	lor	lor	le	le			i	ii	ă<a	a	a	ă
ieți, ieli, iezi, ri	lor	le	i	a	lor	lor	le	le			i	ii	e<a	a	a	ă
ele	lor	le	i	ua	lor	lor	le	le			i	lei		aua	ua	
iri,eri,uni	lor	le	i	a	lor	lor	le	le			i	ii		ea	a	
ări	lor	le	i	a	lor	lor	le	le			i	ii	ă<a	ea	a	
tre, ence, ese	lor	le	i	a	lor	lor	le	le			i	ei	e<ea	a	a	ă
ini,ngi, ei, uri, ti	lor	le	i	a	lor	lor	le	le			i	ii		ea	a	e
ii	lor	le	ei	a	lor	lor	le	le		i	ei	ei		a	a	e
(eț,er,ec,ez,ng)_i, eli	lor	le	i	a	lor	lor	le	le			i	ii	e<ea	a	a	ă
rți, nți	lor	le	i	a	lor	lor	le	le			i	ii	ț<t	ea	a	
oli	lor	le	i	a	lor	lor	le	le			i	ii	o<oa	a	a	ă
izi, nzi	lor	le	i	a	lor	lor	le	le			i	ii	z<d	a	a	ă
ei	lor	le	i	a	lor	lor	le	le			i	i		a	a	e
ăzi	lor	le	le	a	lor	lor	le	le			i	ii	ăz<ad	a	a	ă
ăi	lor	le	i	a	lor	lor	le	le			i	ii	ă<a	lea	a	
țele	lor	le	i	a	lor	lor	le	le			i	elei		ica	a	ă
neregulat: surori																
eți	lor	i	lui	le	lor	lor	i	i	i<e,ț<t		lui	lui		le	le	
ouri, ruri, âuri, euri	lor	le	lui	l	lor	lor	le	le	r<zero	i	lui	ui			l	
oare,oane,oace, oaie	lor	le	ului	ul	lor	lor	le	le	oa<o	e	ului	ui			ul	
uți,ii	lor	le	ului	ul	lor	lor	le	le		i	ului	ui			ul	
ri	lor	le	ului	ul	lor	lor	le	le	r<zero	i	ului	ui			ul	
ere	lor	le	ului	ul	lor	lor	le	le	e<ă	e	ului	ui			ul	
uri	lor	le	ului	ul	lor	lor	le	le	r<zero	i	lui	ui			ul	
ții	lor	le	lui	l	lor	lor	le	le		i	ului	ui			l	
iene	lor	le	ului	ul	lor	lor	le	le	ie<ia	e	ului	ui			ul	
nte	lor	le	ului	ul	lor	lor	le	le	in<ân	e	ului	ui			ul	

TABEL DEFLEXIONARE VERBALĂ- LIMBA ROMÂNĂ, GRUPE DE TERMINAȚII

<u>T1: conjugarea I</u>	
prezent:	imperfect
pers1: nu exista	am
pers2: i	ai
pers3: ă	a
pers4: ăm	am
pers5: ați	ați
pers6: ă	au
perfect simplu	mai-mult-ca-perfect
ai	asem
ași	aseși
ă	ase
arăm	aserăm
arăți	aserăți
ară	aseră

<u>T4: conjugarea I</u>	
prezent:	imperfect
pers1: i	am
pers2: i	ai
pers3: ă	a
pers4: ăm	am
pers5: ați	ați
pers6: ă	au
perfect simplu	mai-mult-ca-perfect
ai	asem
ași	aseși
ă	ase
arăm	aserăm
arăți	aserăți
ară	aseră

<u>T2: conjugarea I</u>	
prezent:	imperfect
pers1: ez	am
pers2: ezi	ai
pers3: ează	a
pers4: ăm	am
pers5: ați	ați
pers6: ează	au
perfect simplu	mai-mult-ca-perfect
ai	asem
ași	aseși
ă	ase
arăm	aserăm
arăți	aserăți
ară	aseră

<u>T5: conjugarea I</u>		<u>Neregulat: a da</u>
prezent:	imperfect	
pers1: dau	dădeam	
pers2: dai	dădeai	
pers3: dă	dădea	
pers4: dăm	dădeam	
pers5: dați	dădeți	
pers6: dau	dădeau	
perfect simplu	mai-mult-ca-perfect	
dădai	dădusem	
dăduși	dăduseși	
dădu	dăduse	
dădurăm	dăduserăm	
dădurăți	dăduserăți	
dădură	dăduseră	

<u>T3: conjugarea I</u>	
prezent:	imperfect
pers1: ez	am
pers2: ezi	ai
pers3: ază	a
pers4: em	am
pers5: ați	ați
pers6: ază	au
perfect simplu	mai-mult-ca-perfect
ai	asem
ași	aseși
e	ase
arăm	aserăm
arăți	aserăți
ară	aseră



<u>T6: conjugarea a 2a</u>	
prezent:	imperfect
pers1: nu exista	eam
pers2: i	eai
pers3: e	ea
pers4: em	eam
pers5: eți	eai
pers6: nu exista	eau
perfect simplu	mai-mult-ca-perfect
ui	usem
uși	useși
u	use
urăm	userăm
urăți	userăți
ură	useră

<u>T7: conjugarea a 2a</u>	
prezent:	imperfect
pers1: ez	eam
pers2: ezi	eai
pers3: ează	ea
pers4: em	eam
pers5: eți	eați
pers6: ează	eau
perfect simplu	mai-mult-ca-perfect
eai	easem
eași	easeși
e	ease
earăm	easerăm
earăți	easerăți
eară	easeră

<u>T10: conjugarea a 2-a <i>Neregulat: a vrea</i></u>	
prezent:	imperfect
pers1: vreau	voiam
pers2: vrei	voiai
pers3: vrea	voia
pers4: vrem	voiam
pers5: vreți	voiați
pers6: vor	voiau
perfect simplu	mai-mult-ca-perfect
vrui	voisem
vruiți	voiseși
vru	voise
vrurăm	voiserăm
vrurăți	voiserăți
vrură	voiseră

<u>T8: conjugarea a 2a</u>	
prezent:	imperfect
pers1: eez	eam
pers2: eezi	eai
pers3: eează	ea
pers4: eăm	eam
pers5: eați	eați
pers6: eează	eau
perfect simplu	mai-mult-ca-perfect
eai	easem
eași	easeși
e	ease
earăm	easerăm
earăți	easerăți
eară	easeră

<u>T11: conjugarea a 2a <i>Neregulat: a avea</i></u>	
prezent:	imperfect
pers1: am	aveam
pers2: ai	aveai
pers3: are	avea
pers4: avem	aveam
pers5: aveți	aveați
pers6: au	aveau
perfect simplu	mai-mult-ca-perfect
avusei	avusesem
avuseși	avuseși
avuse	avusesese
avuserăm	avuseserăm
avuserăți	avuseserăți
avuseră	avuseseră

<u>T9: conjugarea a 2-a</u>	
prezent:	imperfect
pers1: eau	eam
pers2: ei	eai
pers3: ea	ea
pers4: em	eam
pers5: eți	eați
pers6: eau	eau
perfect simplu	mai-mult-ca-perfect
ăui	ăusem
ăuși	ăuseși
ău	ăuse
ăurăm	ăuserăm
ăurăți	ăuserăți
ăură	ăuseră



<u>T12: conjugarea a 3-a</u>	
prezent:	imperfect
pers1: nu exista	eam
pers2: i	eai
pers3: e	ea
pers4: em	eam
pers5: eți	eați
pers6: nu exista	eau
perfect simplu	mai-mult-ca-perfect
sei	sesem
seși	seseși
se	sese
serăm	seserăm
serăți	seserăți
seră	seseră

<u>T13:conjugarea a 3-a</u>	
prezent:	imperfect
pers1: nu exista	eam
pers2: i	eai
pers3: e	ea
pers4: em	eam
pers5: eți	eați
pers6: nu exista	eau
perfect simplu	mai-mult-ca-perfect
ui	usem
uși	useși
u	use
urăm	userăm
urăți	userăți
ură	useră

<u>T16: conjugarea a 4-a</u>	
prezent:	imperfect
pers1: esc	eam
pers2: ești	eai
pers3: ește	ea
pers4: im	eam
pers5: iți	eați
pers6: esc	eau
perfect simplu	mai-mult-ca-perfect
ii	isem
iși	iseși
i	ise
irăm	iserăm
irăți	iserăți
iră	iseră

<u>T14: conjugarea a 3-a</u>	
prezent:	imperfect
pers1: u	am
pers2: i	ai
pers3: e	a
pers4: em	am
pers5: eți	ați
pers6: u	au
perfect simplu	mai-mult-ca-perfect
sei	sesem
seși	seseși
se	sese
serăm	seserăm
serăți	seserăți
seră	seseră

<u>T17: conjugarea a 4-a</u>	
prezent:	imperfect
pers1: nu exista	eam
pers2: i	eai
pers3: e	ea
pers4: im	eam
pers5: iți	eați
pers6: nu exista	eau
perfect simplu	mai-mult-ca-perfect
ii	isem
iși	iseși
i	ise
irăm	iserăm
irăți	iserăți
iră	iseră

<u>T15: conjugarea a 3-a</u>	
prezent:	imperfect
pers1: u	eam
pers2: i	eai
pers3: e	ea
pers4: em	eam
pers5: eți	eați
pers6: u	eau
perfect simplu	mai-mult-ca-perfect
ui	usem
uși	useși
u	use
urăm	userăm
urăți	userăți
ură	useră

<u>T18: conjugarea a 4-a</u>	
prezent:	imperfect
pers1: iesc	iam
pers2: iești	iai
pers3: iește	ia
pers4: im	iam
pers5: iți	iați
pers6: iesc	iau
perfect simplu	mai-mult-ca-perfect
ii	isem
iși	iseși
i	ise
irăm	iserăm
irăți	iserăți
iră	iseră



<u>T19: conjugarea a 4-a</u>	
prezent:	imperfect
pers1: i	iam
pers2: i	iai
pers3: ie	ia
pers4: im	iam
pers5: iți	iați
pers6: ie	iau
perfect simplu	mai-mult-ca-perfect
ii	isem
iși	iseși
i	ise
irăm	iserăm
irăți	iserăți
iră	iseră

<u>T22: conjugarea a 4-a Neregulat: a fi</u>	
prezent:	imperfect
pers1: sunt	eram
pers2: ești	erai
pers3: este	era
pers4: suntem	eram
pers5: sunteți	erați
pers6: sunt	erau
perfect simplu	mai-mult-ca-perfect
fusei	fusesem
fuseși	fuseseseși
fuse	fusesese
fuserăm	fuseserăm
fuserăți	fuseserăți
fuseră	fuseseră

<u>T20: conjugarea a 4-a</u>	
prezent:	imperfect
pers1: nu exista	am
pers2: i	ai
pers3: ă	a
pers4: âm	am
pers5: ați	ați
pers6: ă	au
perfect simplu	mai-mult-ca-perfect
âi	âsem
âși	âseși
â	âse
ârăm	âserăm
ârăți	âserăți
âră	âseră

<u>T23: conjugarea a 4-a</u>	
prezent:	imperfect
pers1: nu exista	eam
pers2: i	eai
pers3: ă	ea
pers4: im	eam
pers5: iți	eați
pers6: ă	eau
perfect simplu	mai-mult-ca-perfect
ii	isem
iși	iseși
i	ise
irăm	iserăm
irăți	iserăți
iră	iseră

<u>T21: conjugarea a 4-a</u>	
prezent:	imperfect
pers1: ăsc	am
pers2: ășți	ai
pers3: ăște	a
pers4: âm	am
pers5: ați	ați
pers6: ăsc	au
perfect simplu	mai-mult-ca-perfect
âi	âsem
âși	âseși
â	âse
ârăm	âserăm
ârăți	âserăți
âră	âseră

<u>T24: conjugarea a 4-a</u>	
prezent:	imperfect
pers1: iu	iam
pers2: ii	iai
pers3: ie	ia
pers4: im	iam
pers5: iți	iați
pers6: iu	iau
perfect simplu	mai-mult-ca-perfect
iui	iusem
iuși	iuseși
iu	iuse
iurăm	iuserăm
iurăți	iuserăți
iură	iuseră

DEFLEXIONARE VERB - LIMBA ROMANA, GRUPE TERMINATII

Verbe: conjugarea I terminatia a

Modul indicativ

conjugarea a II-a terminatia **ea**; conjugarea a III-a terminatia **e**; conjugarea a 4a terminatiile **i, î**

Nr reg	Terminatie:: scad sir	Alt. pers.1	Alt. pers.2	Alt. pers.3	Alt. pers.4	Alt. pers.5	Alt. pers.6	Restrictie: de verificat	adun sir	conjugarea
1	T1	1	1+2	1			1	ăut	a	1
2	T1	1	1+4	1			1	ăsc	a	1
3	T1	1	1	1			1	ătr	a	1
4	T1	6	6+2	6			6	ent	a	1
5	T1		2					lt,et,ânt,it	a	1
6	T2	nu exista						on, ot, or,am,nț	a	1
7	T2	nu exista						nt, u, r,es,ct,rm	a	1
8	T2	nu exista						ul,in,of	a	1
9	T2	nu exista						dr,cet	a	1
10	T2	nu exista						aj	a	1
11	T2	nu exista						i, id	a	1
12	T2	nu exista						iz	a	1
13	T1			3			3	e+(rg,mn, r,st,rs)	a	1
14	T1	5	5	5			5	uc	a	1
15	T1		2	34			34	ăt	a	1
16	T1	35+36	35+36	35+36			35+36	ânc	a	1
17	T1		39	34			34	ăl	a	1
18	T1			34			34	ec,eț	a	1
19	T1	37	37	37			37	tin	a	1
20	T1		2	38			38	pt	a	1
21	T1	40	40	40				ug	a	1
22	T1			38			38	eg	a	1
23	T4							ci	a	1
24	T3							nu	a	1
25	T1	7	7	7			7	ig	a	1
26	T5	nu exista						a da		1
27	T1	nu exista						orice mai putin sit anterioare	a	1
28	T6	1	1	1			1	ăr, ăc	ea	2
29	T6	8+9	8+9	8+9	8+9	8+9	8+9	ed	ea	2
30	T6	9	9	9	9	9	9	șed	ea	2
31	T6	1+9	1+9	1+9	9	9	1+9	ăd	ea	2
32	T7	nu exista						gh	ea	2
33	T6	32	2+32	32			32	ut	ea	2
34	T8	nu exista						cr	ea	2

35	T9	nu exista							b	ea	2
36	T10								Neregulat:	ea	2
									vrea (N1)		2
37	T11								Neregulat:	ea	2
									avea (N2)		2
38	T12	11+21	11+21	11+21	11+21	11+21	11+21		ag, arg	e	3
39	T13	12	2+12	12	12	12	12		at	e	3
40	T12	13	13	13	13	13	13		it	e	3
41	T13	12	12	12	12	12	12		ac	e	3
42	T13	nu exista							ec,ern,ep	e	3
43	T12	14+22	14+22	14+22	14+22	14+22	14+22		oac	e	3
44	T13	12+17	12+17	12+17	12+17	12+17	12+17		așt	e	3
45	T13	17	17	17	17	17	17		eșt	e	3
46	T13	15+17	15+17	15+17	15+17	15+17	15+17		oașt	e	3
47	T12	14+26	14+26	14+26	14+26	14+26	14+26		oarc	e	3
48	T13	16	16	16	16	16	16		oas	e	3
49	T12 mai puțin sit anterioare	24	24	24	24	24	24		c,ng,rn,g	e	3
50	T12	19	19	19	19	19	19		d	e	3
51	T12	nu exista							up	e	3
52	T12	14	14+2	14	14	14	14		oat	e	3
53	T12	20	20+25	20	20	20	20		ân	e	3
54	T13	nu exista							rn	e	3
55	T13		18						rd	e	3
56	T13		25						in	e	3
57	T12		41						un	e	3
58	T14	nu exista							vocala	e	3
59	T15	nu exista							pl	e	3
60	T17	28		28			28		ț	i	4
61	T17		23	30					os	i	4
62	T17			30					orm	i	4
63	T17	nu exista							gh	i	4
64	T17	31	31	31			31		ăr, ăc	i	4
65	T18	nu exista							ă	i	4
66	T19	nu exista							itu,ibu,â	i	4
67	T17	32	32	32			32		ur	i	4
68	T17	6	6+25	6			6		en	i	4
69	T18	nu exista							v,p,s,c,r,l,	i	4
									d,b,t,u,z,n,		4
									j,ver,găr,m		4
70	T17	27		27			27		ș	i	4
71	T22	nu exista							neregulat:	i	4
									fi		4
72	T20			30					or	î	4
73	T23	nu							er	i	4

		exista								
74	T24	nu exista						șt	i	4
75	T21	nu exista						ur	î	4

DEFLEXIONARE VERB - LIMBA ROMÂNĂ, ALTERNANȚE

-

1. a<ă prezent persoanele 1,2,3,6 (eu,tu,el,ei)
2. ț<t prezent persoana 2 (tu)
3. a<e prezent persoanele 3,6 (el,ei)
4. șt<sc prezent persoana 2 (tu)
5. o<u prezent persoanele 1,2,3,6 (eu,tu,el,ei)
6. i<e prezent persoanele 1,2,3,6 (eu,tu,el,ei)
7. gh<g prezent persoanele 1,2,3,6 (eu,tu,el,ei)
8. ă<e prezent 1,6; perfect simplu + mai mult ca perfect 1,2,3,4,5,6
9. z<d prezent 2; pf simplu + mai mult ca perfect 1,2,3,4,5,6
10. o<u prezent 1,6; oa< u prezent 3
11. ă<a imperfect 1,2,3,4,5,6; perfect simplu 1,2 mai mult ca perfect 1,2,3,4,5,6
12. ă<a imperfect, pf simplu, m.m.c.pf 1,2,3,4,5,6
13. ț<t prezent 2, zero< t perfect simplu, m.m.c.pf 1,2,3,4,5,6
14. o<oa prezent 1,2,6, imperfect 1,2,3,4,5,6 pf simplu 1,2; m.m.c.pf 1,2,3,4,5,6
15. o< oa prezent 1,2,6, imperfect 1,2,3,4,5,6 pf simplu si mai mult ca perfect 1,2,3,4,5,6
16. o<oa prezent 1,2,6, imperfect 1,2,3,4,5,6 u<oa pf simplu si m.m.c.pf 1,2,3,4,5,6
17. sc<șt prezent 1,6; pf sp si m.m.c.pf 1,2,3,4,5,6
18. z<d prezent 2
19. z<d prezent 2; zero<d pf simplu, m.m.c.pf

1,2,3,4,5,6

20. a<â pf simplu 1,2,3,4,5,6; ă<â m.m.c pf

1,2,3,4,5,6

21. zero<g pf simplu, m.m.c.pf 1,2,3,4,5,6

22. p<c pf simplu si m.m.c.pf 1,2,3,4,5,6

23. ș<s prezent 2

24. zero< (c, ng, rn, g) pf simplu si m.m.c.pf

1,2,3,4,5,6

25. zero<n prezent 2

26. zero<c perfect simplu si m.m.c.pf

1,2,3,4,5,6

27. s<ș prezent 1,3,6

28. t<ț prezent 1,3,6

29. ă<e prezent 1

30. oa<o prezent 3

31. a<ă prezent 1,2,3,6

32. o<u prezent 1,2,6; oa<u prezent 3

33. i<e prezent 1,2,3,6

34. ă<a prezent 3 sg,pl

35. ă<â 1sg, 2sg, 3sg, pl prez

36. nân < n 1sg, 2sg, 3sg, pl prez

37. a<â 1sg, 2sg, 3 sg, pl prez

38. ea<e 3sg, pl prez

39. e < ă 2 sg, prez

40. o <u 1 sg, 2 sg prez; oa<u 3 sg, pl

41. zero < u 2 sg prez, 1,2,3,4,5,6 pf sp, m.m.c.pf.

Generarea timpurilor verbale corespunzătoare modului indicativ, în limba engleză

Modul indicativ, forma afirmativă: verb_{afirmativ}

Timpul	Clasă verb	Pers.		Altern.	Aux. ₁	Aux. ₂
		3	1,2,4,5,6			
Present Tense Simple	T3, T4, T5	s	-	-		
	T1	es	-	-		
	T2	es	-	y -> i		
Present Tense Continuous	T1, T2, T5	ing	ing	-	Be1	
	T3	[+ cons.] ing ¹⁰⁵	[+ cons.] ing	-	Be1	
	T4	[-e]ing ¹⁰⁶	[-e]ing	-	Be1	
Past Tense Simple	T1, T5	ed	ed	-		
	T4	+d	+d	-		
	T2	ed	ed	y -> i		
	T3	[+ cons.] ed	[+ cons.] ed	-		
Past Tense Continuous	T1, T2, T5	ing	ing	-	Be2	
	T3	[+ cons.] ing	[+ cons.] ing	-	Be2	
	T4	[-e]ing	[-e]ing	-	Be2	
Present Perfect Simple	T1, T5	ed	ed	-	Have1	
Present Perfect Simple	T4	+d	+d	-	Have1	
	T2	ed	ed	y -> i	Have1	
	T3	[+ cons.] ed	[+ cons.] ed	-	Have1	
Present Perfect Continuous	T1, T2, T5	ing	ing	-	Have1	Be3
	T3	[+ cons.] ing ¹⁰⁷	[+ cons.] ing	-	Have1	Be3

¹⁰⁵ Dublează consoana, înainte de a primi terminația specifică aspectului continuu, *ing*.

¹⁰⁶ Elimină fonemul *e*, înainte de a primi terminația *ing*.

¹⁰⁷ Dublează consoana, înainte de a primi terminația specifică aspectului continuu, *ing*.

	T4	[-e]ing ¹⁰⁸	[-e]ing	-	Have1	Be3
Past Perfect Simple	T1, T5	ed	ed	-	Have2	
	T4	+d	+d	-	Have2	
	T2	ed	ed	y -> i	Have2	
	T3	[+ cons.] ed	[+ cons.] ed	-	Have2	
Past Perfect Continuous	T1, T2, T5	ing	ing	-	Have2	Be3
	T3	[+ cons.] ing ¹⁰⁹	[+ cons.] ing	-	Have2	Be3
	T4	[-e]ing	[-e]ing	-	Have2	Be3
Future Tense Simple	T1, T2, T3, T4, T5	-	-	-	1- Shall 1,2,3,4,5,6- Will	
Future Tense Continuous	T1, T2, T5	ing	ing	-	Be4	
	T3	[+ cons.] ing	[+ cons.] ing	-	Be4	
	T4	[-e]ing	[-e]ing	-	Be4	

Modul indicativ, forma negativă: verb_{negativ}

Timpul	Clasă verb	Persoană	Alt. fon.	Aux.1	Aux.2
Present Tense Simple	T1, T2, T3, T4, T5			Do _{1Neg}	
Present Tense Continuous	T1, T2, T5	ing		Be _{1Neg}	
	T3	[+ cons.] ing		Be _{1Neg}	
	T4	[-e]ing		Be _{1Neg}	
Past Tense Simple	T1, T2, T3, T4, T5			Do _{2Neg}	
Past Tense Continuous	T1, T2, T5	ing		Be _{2Neg}	
	T3	[+ cons.] ing		Be _{2Neg}	
	T4	[-e]ing		Be _{2Neg}	
Present Perfect Simple	T1, T5	ed		Have _{1Neg}	
	T4	+d		Have _{1Neg}	

¹⁰⁸ Elimină fonemul *e*, înainte de a primi terminația *ing*.

¹⁰⁹ Dublează consoana, înainte de a primi terminația specifică aspectului continuu, *ing*.

	T2	ed	y -> i	Have _{1Neg}	
	T3	[+ cons.] ed		Have _{1Neg}	
Present Perfect Continuous	T1, T2, T5	ing		Have _{1Neg}	Be ₃
	T3	[+ cons.] ing		Have _{1Neg}	Be ₃
	T4	[-e]ing		Have _{1Neg}	Be ₃
Past Perfect Simple	T1, T5	ed		Have _{2Neg}	
	T4	+d		Have _{2Neg}	
	T2	ed	y -> i	Have _{2Neg}	
	T3	[+ cons.] ed		Have _{2Neg}	
Past Perfect Continuous	T1, T2, T5	ing		Have _{2Neg}	Be ₃
	T3	[+ cons.] ing		Have _{2Neg}	Be ₃
	T4	[-e] ing		Have _{2Neg}	Be ₃
Future Tense Simple	T1, T2, T3, T4, T5			1, 4- Shall not / shan't 1,2,3,4,5,6- Will not / won't	
Future Tense Continuous	T1, T2, T5	ing		Be _{4Neg}	
	T3	[+ cons.] ing		Be _{4Neg}	
	T4	[-e]ing		Be _{4Neg}	