

REZOLVAREA COMPUTERIZATĂ A PROBLEMELOR DE ORDIN LINGVIC

Valentin CIJACOVSKI, prof. univ., dr. hab.
Universitatea Liberă Internațională din Moldova
Vladimir PROCOPCIUC, programator
Universitatea Tehnică a Moldovei

Experimentul propus de traducerea automată a scrisorilor de corespondență de afaceri engleză în limba română este dedicat simulării procesului de analiză și sinteză morfologico-sintactic-semantică realizat în creierul uman la generarea și perceperea unui text specializat și formatizat în corespundere cu cerințele computerului prin algoritimizare și programare.

Cuvinte-cheie: *post-redacție, experiment, traducere automată, text specializat.*

The present experiment of computer-aided business letters' translation from English into Romanian is devoted to the simulation of the process of morphologic, syntactic and semantic analysis and synthesis taken place in the human mind on specialized text's comprehending and generating according to the computer requirements of algorithm creation and programming.

Key-words: *post-editing, experiment, automatic translation, specialized text.*

Drept premise pentru acest studiu au servit:

- a) modelele reproductiv ingineresti lingvistice (MRIL), capabile de a funcționa în automatul cibernetic-lingvistic (ACL);
- b) prezența condițiilor informativ-lingvistice care au favorizat utilizarea manifestărilor mecanismului sinergetic-lingvistic propriu limbii naturale (LN) și activității verbal-cognitive umane (AVCU) [2, 3].

Primele fac posibile o reprezentare modelată prin analogie a obiectelor reale, adică prin modele lingvistice formalizate care pot fi folosite de ACL (computer) la rezolvarea problemelor lingvistice prin programarea lor în procesul traducerii automate.

Cele de a doilea tip ne permit să ținem cont de manifestările sinergetice de reproducere și autoreglare ale energiei interioare a LN și AVCU care stau la baza algoritimizării procesului de prelucrare automată a datelor de cunoștință și celor informativ-lingvistice în cursul procesului de traducere.

Alegerea premiselor sus-menționate a fost predeterminată de unele idei, chemate să asigure legătura între partea formală și cea naturală a acestui proces de prelucrare automată:

- 1) textul specializat supus prelucrării trebuie să aparțină la un anumit domeniu de aplicație (DA);
- 2) procesul de prelucrare automată trebuie să fie consecutiv subdivizat în așa mod ca, pe de o parte, să ne putem folosi de rezultatele subdiviziunii precedente în cea ulterioară, și, pe de altă parte, să favorizăm manifestările mecanismului sinergetic lingvistic într-o continuitate neîntreruptă;
- 3) coordonarea modului de codificare a elementelor lingvistice, participante a procesului de traducere automată; el trebuie să reflecte caracterul consecutiv și neîntrerupt al operațiunilor de realizat;
- 4) analiza și sinteza părților de vorbire folosite în text trebuie să fie realizată într-o continuitate de conținut la nivel individual (cuvânt) și colectiv (segment și propoziție).

Diviziunile principale ale procesului de traducere automată au fost definite ca segmentare, schematizare și traducere propriu zisă. La rândul lor, ele presupun subdiviziunile necesare. Ca material experimental vom utiliza scrisorile de afaceri engleze [1].

Problema tehnică care o rezolvă invenția dată constă în elaborarea condițiilor care ar favoriza reflectarea legilor interioare ale limbii (sinergia lingvistică).

Segmentarea automată a propoziției engleze se realizează pe baza sistemului de codificare a părților de vorbire engleze, a setului de reguli pentru eliminarea dualității de interpretare și tabelii de indicii; schematizare, adică atribuirea automată segmentelor de propoziție deja detectate de automatul cibernetic lingvistic (ACL) funcțiilor lor sintactico-semantice concomitent cu fixarea grafică a ordinii lor de succesiune, și deambiguizarea segmentelor, adică eliminarea automată a polisemiei propriie majorității părților de vorbire engleze, pe baza setului de reguli de dezambiguizare programată, recodarea segmentului englez la cel român în cazul lipsei de coincidență între formele de exprimare engleze corespunzătoare celor române, stabilirea ordinii de cuvinte în segmentul român, filtrarea, adică alegerea echivalentului român de traducere a cuvântului corespunzător englez, fixarea caracteristicilor morfologice și sintactice constante ale părților de vorbire a segmentului român analizat, stabilirea și fixarea de valență generală exterioară (LDVE), stabilirea și fixarea legăturilor de valență generală interioară (LDVI), fixarea finală a caracteristicilor morfologice și sintactice inclusiv cele neconstante, alegerea automată a variantei de

traducere (W), adică fixarea paradigmelor de declinare și conjugare a formei cuvântului cerută de segmentul în care este localizat; fixarea finală a rezultatului operațiilor de traducere din limba engleză în cea română.

Rezultatul tehnic constă în realizarea unei traduceri automate a scrisorilor comerciale engleze în limba română, fără post-redactare. Mai departe vom prezenta exemple de realizare a metodei de traducere automată fără post-redacție.

În partea ce urmează vor fi descrise în măsură corespunzătoare ordinea de prelucrare automată a textului/propoziției specializate engleze, adică consecutivitatea și conținutul elementelor acestui proces continuu care asigură trecerea la realizarea traducerii în limba română [4, 5, 6].

Aici se includ următoarele: segmentarea, precedată de eliminarea dualităților de interpretare a unor părți de vorbire, schematizarea și dezambiguizarea. Toate aceste operațiuni presupun prezența codificării părților de vorbire engleze, a seturilor de reguli care reglementează realizarea lor la nivel lingvistic și cel de programare.

În calitate de material de bază ne-am folosit de manualul *Corespondența de afaceri în limbile române și engleză* (Editura Teora) [1].

Sub noțiunea de segmentare se înțelege fixarea automată a părților componente de conținut ale propoziției analizate care participă la formarea sensului ei. Segmentul poate fi prezentat printr-o parte de vorbire luată aparte (conjuncția *while*) sau printr-o îmbinare de cuvinte (*your proposal*) (v. tab. 3).

Determinării automate a segmentelor precede elaborarea modelelor reproductive ingineresti (MRI) care simulează ordinea de generare și percepere a propoziției analizate dictată de mecanismul sinergetic lingvistic propriu LN și ACVU. Aceste modele sunt bazate pe utilizarea proprietăților morfologice și sintactico-semantice a părților de vorbire engleze care se manifestă diferit în dependență de rolul jucat de ele la formarea sensului exprimat de propoziția în cauză.

Reprezentarea lor sub forma tabelelor de indicii codificați ajustată de setul de reguli corespunzător asigură funcționarea lor automată.

În acele cazuri când este necesară eliminarea interpretării a unei părți de vorbire echivoce înainte de a realiza operațiunea de segmentare automată a treia literă a codului A va fi înlocuită prin R.

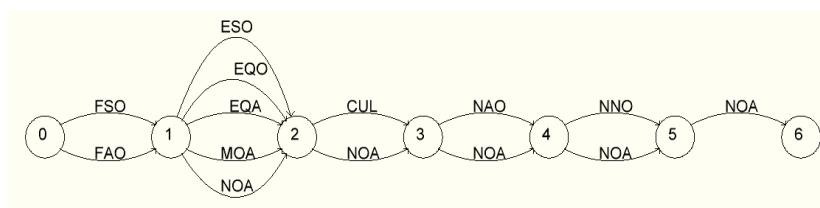
De exemplu: prepoziția *to* se marchează cu codul PTR, dat fiind că ea poate introduce infinitivul sau un segment prepozițional. În primul

caz R va fi înlocuit prin O (PTO), în cazul al doilea PTR va fi înlocuit prin PSO (prepoziția simplă neambiguă). Vezi ultima propoziție a scrisorii, unde ambele cazuri sunt rezolvate: *to meet* și *to your prompt reply* (v. scrisoarea).

Conjuncția participială *considering that*, fiind detectată de calculator prin primul ei membru *considering* marcat cu codul VGR, se supune operațiunilor dictate de regula corespunzătoare. Ultima prevede detectarea următoarei părți de vorbire *that* (CTA) și în cazul prezenței sale permite fixarea segmentului conjuncțional marcat CPO prin două intervale (_ _).

Tabele de indicii prezintă în sine o sistematizare de coduri aranjate în coloanele corespunzătoare (1, 2, 3,...) a părților de vorbire engleze care participă la formarea segmentelor propozițiilor analizate și numerotate de la 0 până la 9 corespunzător numărului părților de vorbire engleze codate. Când este necesar, adăugăm litere la cifrele utilizate pentru a lărgi diapazonul lor de includere (pentru pronume noi introducem 4A, 4B, 4C). Ca exemplu vom folosi tabelul de indicii jos menționat.

nr. col.	1	2	3	4	5	6	7	8	9	10
Codurile părților de vorbire	FSO FAO	ESO EQO EQA MOA NOA	CUL NOA	NAO NOA	NNO NOA	NOA	-	-	-	-



Acest tabel se referă la toate segmentele introduse prin pronumele posesiv. El este codat prin FSO atunci când el este utilizat ca pronume adjectival (forma asociată) și prin FAO când este utilizat ca un substantiv (formă absolută). Ambele coduri ocupă prima coloană. În următoarele coloane sunt indicate codurile părților de vorbire engleze care pot participa la formarea segmentelor introduse de FSO și FAO (prin E se introduce adjectivul, prin M – numeralul, prin N – substantivul, prin C

– conjuncția).

Să presupunem că calculatorul începe analiza unui segment introdus prin pronumele posesiv (FSO) (v. tabelul de indici nr. 4B). Marcând prezența sa în prima coloană a tabelului, calculatorul fixează prezența sa printr-un interval (*your_*) și trece la analiza cuvântului următor – *proposal* – codat cu NOA. Pentru a hotărî apartenența sa la segment este necesară adresarea la coloana a doua pentru confirmarea prezenței sale. Răspunsul fiind afirmativ, calculatorul îl fixează printr-un interval și trece la analiza cuvântului următor al propoziției – *for* (PSO). Se repetă adresarea la tabelul de indicii, dar de data aceasta la coloana a treia. Absența PSO în coloană trebuie să fie interpretată de calculator ca un început de segment nou care trebuie să fie despărțit de cel precedent prin două intervale (*_ _*) – *your_proposal_ _*.

În același mod a fost fixat sfârșitul primului segment al propoziției analizate introdus prin *we* (FNO), dar cu adrese la un alt tabel de indiciu.

Pentru punerea în funcțiune a operațiilor de codificare automată ne-am folosit de 16 tabele de indicii: de 5 pentru segmentele introduse de verb, de 3 pentru segmentele introduse de pronume și de câte una pentru restul părților de vorbire.

Următoarea etapă de prelucrare automată a propoziției engleza are ca scop de a atribui segmentelor detectate de ACL funcțiile lor sintactico-semantice în propoziție și de a fixa concomitent grafic oridnea lor de succesiune.

Am subliniat deja că părțile de vorbire purtătoare de informație principale sunt substantivul și verbul. În textele de specialitate, în scrisorile de afaceri, concomitent cu ele sunt legate corespunzător fenomenele semantice de temă și remă [4]. Segmentarea a demonstrat că aceste părți de vorbire sunt componentele principale ale segmentelor detectate și că pe baza lor se formează 3 din cele 5 tipuri de segmente determinate de noi. În ordinea lor de prezentare în tabel ele sunt următoarele:

- 1) segmentul conjuncțional-prepozițional, codat cu zero (SCP0);
- 2) segmentul verbal-rematic, codat cu 1 (SVN1), care îmbină funcția sintactică de predicatal propoziției analizate cu cea semantică de purtător de remă, adică de informație relatantăla nouitate;
- 3) segmentul nominal-tematic, codat cu 2 (SNT2), care îmbină rolul de subiect al propoziției analizate cu cel de purtător de informație cunoscută ambelor părți de comunicare;

4) segmentul propozițional-nominal-determinativ, codat cu 4 (SAD4), care fiind introdus de o prepoziție prezintă informație adăugătoare la conținutul informației nominal-tematice;

5) segmentul adverbial-determinativ, codat cu 4 (SAD4), care precizează unele circumstanțe legate de segmentele la care el se referă.

Informația codată 0, 1, 2, 3, 4, se atribuie segmentului urmînd operațiunile de segmentare: 0 – celui conjuncțional / semn de punctuație; 1 – celui unde verbul predicativ are rolul principal; 2 – celui unde substantivul determină conținutul; 3 – celui introdus de prepoziție referită la un substantiv și 4 – celui care este introdus de un verb fără prezența verbului sau substantivului (v. tab. 3).

Reguli de schematizare:

1. CPS (Semnele de punctuație și conjuncțiile)

CSO + ...

CAO + ...

CPO + CTO + ...

CTO + ...

CUC + ...

IOC + ... → [0]

POC + ...

CUL + ...

CAL + ...

CAC + ...

2. SVR (Verbal-rematic)

... + [VAO] + ...

... + [VIO] + ...

... + [VNO] + ...

... + [VMO] + ...

... + [VLO] + ... → [1]

... + [VSO] + ...

... + [VBO] + ...

... + [VGO] + ...

3. SNT (Nominal-tematic)

... + [NOO] + ... → [2]

... + [FAO] + ...

4. SPND (Prepozițional-nominal-determinativ)

PSO + ... → [3]

5. SAD (Adverbial-determinativ)

RAO + ... + [=/NOO] + ...
RAO + ... + [=/ZOO] + ...
RAO + ... + [=/VAO] + ...
RAO + ... + [=/VIO] + ...
RAO + ... + [=/VNO] + ... → [4]
RAO + ... + [=/VMO] + ...
RAO + ... + [=/VLO] + ...
RAO + ... + [=/VSO] + ...
RAO + ... + [=/VBO] + ...
RAO + ... + [=/VGO] + ...

Codificarea primară a părților de vorbire engleze a servit ca bază morfologic-sintactică pentru segmentarea textului scrisorilor de afaceri engleze. Următorul pas pentru asigurarea unei traduceri automate a acestor segmente organizate în componența propozițiilor fără a se adresa la un post-redactor va fi realizat în cursul operațiunilor de dezambiguizare a acestora. Sensul acestor operațiuni automate constă în eliminarea poli-semiei propriie majorității părților de vorbire engleze, fără de care nu este posibilă realizarea problemei sus-menționate.

Segmentarea propoziției engleze, determinînd componenții sintactico-semantic, creează premisele realizării unei dezambiguizări corespunzătoare a părților de vorbire engleze care participă la formarea segmentelor și pe baza cărora se formează un set de reguli de dezambiguizare.

În calitate de prima premisă poate fi determinată o anumită selecționare a componenților de segment. Ea presupune o astfel de îmbinare a diferitor părți de vorbire care ar asigura redarea sensului corespunzător și prin urmare determină variant de dezambiguare a părții de vorbire ambigue. Vom compara segmentele *We have dully received* și *We have also to point out*. La segmentare cuvintele subliniate vor fi codate prin VAO și VMO, adică ca auxiliar și modal. Toate aceste posibile acțiuni se fixează prin reguli.

În calitate de a doua premisă, putem considera faptul că segmentele sunt formate în funcție de destinația lor sintactico-semantică ca verbal-rematice, nominal-tematice, prepozițional-adverbial-determinative și conjuncționale. Ca urmare fiind diferite după conținut, selecționarea membrilor lor componenți la fel este diferită.

Adresându-ne la scrisoarea procesată, putem diferenția 45 de segmente: 13 verbal-rematice, 9 nominal-tematice, 15 prepozițional-adverbial-determinative și 8 conjuncționale.

În fiecare din aceste segmente cel puțin un membru-component conține A în poziția a treia a codului său.

Înainte de a trece la dezambiguizare este necesar ca pentru fiecare parte de vorbire să fie determinată variantele de ambiguitate. De exemplu, pentru *have*, codul VAA, sunt prevăzute patru forme posibile dezambiguate: VAO – verb cu funcția de auxiliar; VNO – verb cu funcția de fictiv; VMO – verb cu funcția de modal și VJO – verb cu funcția de infinitiv.

Analizând trecerile posibile a semnelui de ambiguitate A la O, putem afirma că în majoritatea cazurilor (~80%) O înlocuiește pe A în prima variantă fixată de vocabular pentru fiecare cuvânt. Acest fenomen ne permite de a realiza procesul de dezambiguizare fără a aplica reguli speciale (v. reguli de dezambiguizare).

Reguli de dezambiguizare:

- 1) (VAA) + (PTO) → (VMO) + PTO
- 2) (VAA) + RAO + PTO → (VMO) + RAO + PTO
- 3) FNO + (VAA) → FNO + (VNO)
- 4) (NOA) + NOA → (NEO) + NOO
- 5) FNO + (VAA) + [=/VDA] → FNO + (VNO) + [=/VDA]
- 6) VIA + (RAA) → VIA + (RVO)
- 7) FNO + (VIA) → FNO + (VNO)
- 8) (VEA) + EQO → (VLO) + EQO
- 9) (VEA) + RAO + EQO → (VLO) + RAO + EQO
- 10) (VEA) + ANO + RLO + EQO → (VLO) + ANQ + RLO + EQO
- 11) restul: (**A) → (**O)

Traducerea scrisorii de afaceri engleze în limba română

Procesul de traducere a textului scrisorii de afaceri engleze prevede un șir de operațiuni automate pe bază de reguli programate care se realizează în mod consecutiv în limitele de intrasegmentare. Aici sunt prevăzute două etape:

1. Etapa de transfer, adică de trecere de la sistemul de atribuire a codurilor părților de vorbire engleze la cele române, și de la ordinea de cuvinte segmentară engleză la cea română; ea se încheie cu fil-

trarea și fixarea echivalentului de traducere urmat de fixarea caracteristicilor morfologice și sintactice *constante* a membrilor componenți ai segmentului analizat.

2. Etapa propriu zisă de traducere cărei îi precede detectarea legăturilor determinante de valență la nivelul intersegmentar și cel intrasegmentar. Pe baza ultimelor date se completează șirul de caracteristici morfologice și sintactice ale părților de vorbire române necesare la determinarea automată a echivalentelor de traducere în forma cerută și corespunzătoare sensului exprimat prin ele. Echivalentul român fiind fixat, computerul prezintă utilizatorului traducerea finală. Ultima nu necesită o post-redactare.

Această operațiune se impune mulțumită faptului că nu există coincidență deplină între formele de exprimare a sensului englez în cel corespunzător român la nivelul părților de vorbire. De exemplu, dacă nu vom adresa la funcția morfologic-sintactică a articolului ca parte de vorbire utilizată în ambele limbi, vom constata nu numai o diferențiere calitativă dar și una cantitativă. Articolul hotărât român se alipește la sfârșitul substantivului, alcătuind o grupare morfemică unitară, în care componentele: radical-desinență + articol (*preț-u-l*) apar într-o ordine fixă și nedisociabilă (în engleză – *the price*). În limba română există în plus la cea engleză articolul posesiv sau genetival și articolul demonstrative sau adjectival.

Rezultatele segmentării și schematizării automate a scrisorii de afaceri engleze (în acord cu caracterul informative și poziția în propoziție)

Tabelul nr. 3

Nr. seg.	Tipurile de segmente				
	0. semnele de punctuație și conjuncțiile – CPS	1. Verbal-rematic – SVR	2. Nominal-tematic – SNT	3. Prepozițional-nominal-determinativ – SPND	4. Adverbial-determinativ – SAD
1		We have duly received			
2			your proposal		
3				for the supply	
4				of textile	

5				With pattern- book and price-lists	
6					
1	While				
2		apreciating			
3			the good quality		
4				of all your samples	
5					
6		We find			
7			prices		
8				of patterns and 236	rather high
9					
10.				for our market.	
1		We have also to point out			
2	that				
3			very good similar materials		
4		are now available			
5				From other sources	
6				At prices	
7				from 7	
8				to 10 %	
9					below yours.
1	As				
2		we have			
3				in mind	
4			an order		
5				of some 500 yards	
6				of each sample	
7					

8		we would like to know			
9	If				
10		could give us			
11			a slightly better price		
12					at least
13				for the two above-mentioned patterns	
14					
15	Considering that				
16			our markets		
17		is highly competitive			
18					
19			a keen price		
20		is essential			
1		We trust			
2	that				
3		you will be willing to meet			
4			Our terms		
5	And				
6		look forward			
7				to your promptly reply.	

Reguli de recodare:

- 1) ... + PTO + VIO + [RVO] + ... → ... + CSO + VNO + ...
- 2) ... + VIO + RVO + ... → ... + VNO + ...
- 3) ... + AHO + MCO + ... → ... + ADO + MCO + ...

- 4) ... + AHO + [=/MCO] + ... → ... + [=/MCO] + ...
- 5) ... [=/PSO] + FAO + ... → ... + [=/PSO] + APO + FSO + ...
- 6) ... + VGO → VNO + ...
- 7) PSO + NOO (at least) → RAO – frazeologism
- 8) ... + FIO + ZOO + ... → ... + RAO + Z?OO + ...
- 9) ... + EQO + EQO + ... → ... + EQO1 + EQO2 + ...
- 10) ... + CPO + CTO + ... → ... + CPO + ...
- 11) ... + PSO + FIO + ... → ... + APO + FIO + ...
- 12) ... + PSO + FAO + ... → ... + RFO + APO + FSO + ...
- 13) ... + FAO → ... + APO + FSO

Spre diferențiere de segmentul corespunzător englez ordinea de cuvinte a părților de vorbire componente a segmentului român este dictată nu numai ca rezultat al recodării unor părți de vorbire engleze mulțumită schimbărilor corespunzătoare morfologic-sintactice, dar ca urmare a unor impuneri de ordin semantic exprimate prin respectarea sensului exprimat de acestea.

Ca regulă pot fi evidențiate următoarele succesiuni a părților de vorbire :

- 1) substantivul (NOO) este urmat de adjectiv(EQO/ECO);
- 2) substantivul adjectival (NEO) urmează substantivul neadjectival (NOO);
- 3) ordinea de succesiune a două adjective (EQO1 + EQO2) este inversă (EQO2 și EQO1);
- 4) substantivul (NOO) este urmat de pronumele posesiv (PSO);
- 5) pronumele personal (FNO) nu este evidențiat în segmentul verbal-rematic (SVR), el păstrează influența sa (persoana și numărul) în formă verbului de conjugat;
- 6) adverbul (RAA/RLO) nu poate fi despărțit de adjectivul (EQO/ECO) pe care-l determină și poate despărți forma compusă a verbului de conjugat – urmează după ea;
- 7) pronumele indefinit, care poate realize diferite funcții sintactic-tematice ca cele de pronume-nume sau uneori de adverb, întotdeauna precede (NOO) sau grupa sa.

Setul de reguli care soluționează problemele ordinii de cuvinte a segmentului român.

Reguli de ordonare:

- 1) FNO+... +... → ... + ...

- 2) ... + FOO → + ...
- 3) ... + RLO + VDO + ... → ... + VDO + RLO + ...
- 4) ... + NEO + NOO + ... → ... + NOO + NEO + ...
- 5) a. ... + FSO + NOO + ... → ... + NOO + FSO + ...
- b. ... + EQO + NOO + ... → ... + NOO + EQO + ...
- c. ... + FIO + FSO + NOO + ... → ... + FIO + NOO + FSO + ...
- d. ... + RAO + EQO1 + WQO2 + NOO + ... → ... + NOO + EQO2
 + RAO + EQO1 + ...
- e. ... + RLO + ECO + NOO + ... → ... + NOO + RLO + ECO + ...
- f. ... + RLO + EQO + NOO + ... → ... + NOO + RLO + EQO + ...
- g. ... + FSO + EQO + NOO + ... → ... + NOO + FSO + EQO + ...

Una din problemele care premerg realizării unei traduceri automate fără post-redactor este condiționată de elaborarea sistemului de dezambiguire la nivelul semantic. În primul rând este vorba de termeni.

Sistemele propuse de noi reiese din faptul că fiecare cuvânt poate fi interpretat ca termen cu condiția că el este membru al unei sisteme terminologice lexical-semantice. Ocupând un anumit loc în această sistemă el devine monosemantic. De exemplu, termenul englez *proposal* cu uzuală interpretare *propunere* poate fi tradus ca *ofertă*. Totul depinde, cum am afirmat mai sus, de locul pe care-l ocupă în sistem și domeniul de aplicație „Scrisori de afaceri engleze”, corespunzător – „Correspondență de afaceri română”. Luând în considerație că acest domeniu de aplicație include mai mult de douăzeci de subdomenii de aplicație, care la rândul lor se subdivizează în microcâmpuri de ordinul 1, 2, 3, etc. devine posibilă utilizarea noțiunii de *proposal* cu semnificațiile *propunere* și *ofertă* în anumite cazuri, adică în dependență de locul ocupat de acest termen în sistemul lexical-semantic terminologic „Correspondența de afaceri română”.

Când acest termen este fixat în subdomeniul de aplicație „comanda” într-o scrisoare de răspuns la o ofertă, atunci el poate fi interpretat ca *ofertă*, în celelalte cazuri ca *propunere*.

Procesul de filtrare a echivalentului român început după stabilirea ordinii de cuvinte înainte de determinarea caracteristicilor constante ale membrilor de segment va continua și în decursul soluționării problemelor legate de fenomenele legăturilor determinante de valență exterioară (LDVE) și celor de valență interioară (LDVI).

Fiecare parte de vorbire română, în primul rând cele flexibile, deve-

nind parte componentă a segmentului (v. mai sus stabilirea ordinii de cuvinte) aduce cu sine aportul său la formarea sensului acestui segment. Acest aport este marcat prin anumite caracteristici morfologic-semantice care permit participarea la formarea componentelor segmentului. Aceste caracteristici sunt numite de noi *constante* sunt reflectate prin vocabular și precizate în cursul procesului de dezambiguare. Pentru substantiv ele sunt *genul* și *numărul*, iar pentru verb – *verb noțional*, *auxiliar*, *modal*, *semimodal*, *infinitiv*, *gerunziu*, *participiu*. Ultimele caracteristici participă la formarea diferitor forme ale verbului de conjugat. Celelalte părți de vorbire li se atribuie de asemenea caracteristicile construite corespunzător.

Mesajul generat de autor în forma sa expresivă induce diferite părți de vorbire ale limbii date care fiind îmbinate formează propoziții supuse segmentării, adică relevării părților lor componente de conținut. Pentru formularea adecvată a acestui conținut părțile de vorbire componente necesită o coordonare a formelor lor de expresie ceea ce devine posibil numai atunci când sunt respectate regulile gramaticale de acord, recțiune și valență generalizate sub denumirea de legături determinante de valență exterioară (LDVE) sau interioară (LDVI). Ultima diferențiere de denumire se bazează pe existentele legături intersegmentare și intra-segmentare.

Pentru a asigura realizarea procesului de traducere automată fără post-redactor devine necesar a pune la dispoziția computerului toată informația privind nu numai caracteristicile constante a părților de vorbire implicate în formarea segmentelor, dar și caracteristicile lor neconstante care ar exclude orice deviere de la forma de expresie a părții de vorbire cerută în situația concretă de membru al unui segment, adică al unei unități sintactic-semantice de conținut. De exemplu, pentru substantiv în afară de gen și număr este necesar să fie determinate modul de articulare (cu articolul hotărât sau nehotărât) și cazul: *for the supply* – *pentru livrarea* (subst., fem., sing., hot., at.). Vom desena în trăsăturile generale modul de realizare a LDVE și LDVI.

Sensul general al acestei operațiuni este de a determina modul de funcționare ale acestor legături intersegmentare, pe baza cărora se realizează precizarea formelor de expresie a acelor părți de vorbire care sunt supuse regulilor corespunzătoare de acord, recțiune și valență.

Manifestarea acestor legături poate să fie exprimată ca regulă sub două aspecte:

- a) aspectul de poziționare a unui segment față de altul;
- b) aspectul de influență a unui din membrii segmentului precedent asupra unuia din membrii segmentului care urmează.

În primul caz este vorba de ordinea de succesiune a SVR și SNT prin care se determină cazul nominativ sau acuzativ al substantivului sau grupei nominale a SNT. El este nominativ dacă SNT precede SVT și invers. Dat fiind în limba română formele pentru aceste cazuri coincid ele sunt luate în cont numai formal. În cazul al doilea influența unuia din membrii segmentului precedent se manifestă diferit în dependență de tipul segmentului următor:

- 1) dacă acest segment este introdus de prepoziție (SPND), atunci influenței se supune prepoziția care-l introduce: se determină varianta de traducere și corespunzător cazul. De exemplu, când segmentul *your proposal (oferta Dvs.)* precede segmentul *for the supply (pentru livrarea)*, atunci *proposal* determină varianta de traducere *for* și corespunzător, acuzativul;
- 2) dacă segmentul este un segment verbal introdus de verbul auxiliar (VAO), atunci genul, numărul, forma hotărâtă sau nehotărâtă și cazul sunt determinate de substantivul / grupa substantivală a segmentului nominal tematic (SNT) care-l precede. De exemplu, în segmentul *are now available (sunt acum accesibile)* genul, numărul, forma hotărât sau nehotărât și cazul adjectivului predicativ (*accesibil*) vor fi determinate de substantivul *materials* (material) din segmentul precedent *very good familiar materials (materiale obișnuite foarte bune)*;
- 3) dacă acest segment prezintă cu sine o formă personală a verbului fără prezența pronumelui personal corespunzător, atunci detectarea acestui pronume devine o necesitate. De pildă, acest pronume personal introduce prima propoziție a unui complex de propoziții legate cu prima prin coordonare sau subordonare. Propoziția care este introdusă prin conjuncția *and (și)* prezintă cu sine segmentul verbal lipsit de pronumele personal. Rămâne ca acestui segment să fie readresat printr-o regulă LDVE numărul și persoana numită de pronumele personal al primei propoziții. De exemplu, în propoziția complexă: *We trust that you will be willing to meet our terms and look forward to your prompt reply (Sperăm că veți fi dispuși să satisfaceți condițiile noastre și așteptăm cu nerăbdare răspunsul Dvs. prompt)* pronumele personal *We* va indica mulțumit prezenței

and persoana și numărul verbului de conjugat *look forward* (1 plural) – se întâlnesc și alte deosebiri de realizare a aspectelor sus-numite soluționarea tuturor cazurilor LDVE ne reglementează pe baza regulilor elaborate de noi.

Principala sarcină a operațiunilor LDVI este de a completa caracteristicile constante ale părților de vorbire flexibile participante la generarea segmentelor cu cele neconstante, adică cu cele care pot fi determinate numai în context (în propoziția scrisorii de afaceri engleze), parțial pe baza LDVE (intersegmentării și complet pe baza LDVI (intra-segmentară).

Operațiunile de schematizare ne permit repartizarea segmentelor propoziționale corespunzător funcțiilor lor sintactico-semantic, ceea ce face posibil în interesul elaborării regulilor LDVI de a le subdiviza în segmentele grupei verbale (segmentele verbal-rematice), segmentele grupei nominale (segmentele nominal tematice și segmentele prepozițional nominal determinative) și segmentele auxiliare (conjuncționale și adverbial-determinative) (v. Tabelul 4). Primele două (segmentele verbal-rematice și nominal-tematice) vor servi ca bază pentru soluționarea tuturor problemelor legate de LDVI.

Ambele tipuri de segmente determină într-un mod diferit atribuirea caracteristicilor morfologic-sintactice părților de vorbire flexibile participante la formarea acestor segmente.

Segmentul verbal rematic (SVR) poate fi subdivizat în două subtipuri: pur verbal și verbal-predicativ. Cel pur verbal necesită atribuirea acelor caracteristici care vor permite determinarea modului, timpului, persoanei și numărului. În ceea ce privește cel verbal-predicativ, aici caracteristicile verbului de conjugat vor fi completate atât pe baza LDVE atât și în cursul realizării regulilor LDVI.

Ca exemplu propunem două segmente: (*We*) *have duly received* – (*we*) *have received duly* – *am primit în timp și is essential* – *este esențial*.

În primul segment îmbinarea *have received* (VAO_VDO) permite computerului a preciza forma verbului de conjugat: indicativ, perfectul compus (*am primit*). În cel de-al doilea segment partea verbală *is* nu cere adresarea la segmentul precedent dat fiind forma *is* are caracteristicile constante de verb noțional, prezent indicativ, persoana 3, singular (*este*), iar partea predicativă (*essential*) pe baza LDVE obține din segmentul precedent (*a keen price* – *un preț avantajos*) caracteristi-

cile de neutru, singular, nominativ – esențial.

Segmentele grupei nominale în general se caracterizează prin rolul dominant al substantivului introdus în segmentul corespunzător cu denumirea genului și numărului. Rămâne de a obține informația prin care vor fi determinate forma hotărâtă sau nehotărâtă și cazul.

Forma hotărâtă este condiționată prin prezența în segment a articolului hotărât (*for the supply – pentru livrarea*), prin prezența pronumelui posesiv (*your proposal – propunerea Dvs.*), prin prezența pronumelui indefinit **all** (*of all your samples – celor tuturor mostrelor Dvs.*), prin prezența numeralului cifrat după substantiv (*of patterns 235 – la eșantioanele 235*); absența părților de vorbire sus-menționate înseamnă prezența formei nehotărâte a substantivului.

În ceea ce privește informația referitor la concretizarea cazului (el se referă la segmentele nominale introduse de prepoziție) și punctul *a* al aceluiași subparagraf care se ia în cont numai formal.

Un caz deosebit reprezintă segmentul introdus de prepoziția *of*, când aceasta fiind recodată în articol posesiv atribuie substantivului cazul genitiv (*of all your samples – ale tuturor mostrelor Dvs.*).

Părțile flexibile ale segmentului nominal coordonează caracteristicile neconstante cu cele ale substantivului, de exemplu, în segmentul *very good similar materials – materiale similare foarte bune*, adjectivelor *similar* și *bune* sunt atribuite următoarele caracteristici (ele coincid): adjectiv, atributiv, neutru, plural nehotărât și nominativ.

Setul de reguli, care prevede modul de rezolvare a tuturor cazurilor legate de problema LDVE este indicat în continuare.

Seturi de reguli referitoare la legăturile determinate de valență exterioară (LDVE) și legăturile determinate de valență interioară (LDVI), intersegmentară și intrasegmentară.

La nivelul segmentelor Verbal-Rematice SVR:

LDVI

- 1) Dacă FNO (we, you) introduce SVR, atunci el nu se traduce, dar determină persoana și numărul VAO (verbului de conjugat);
- 2) VAO+VDO – perfectul compus;
- 3) Dacă VGO formează SVR, el se transformă în NOO și se traduce prin prezent indicativ;
- 4) Dacă VIO urmează după FNO, atunci el devine VNO;

- 5) Dacă PTO + VIO urmează după VMO, atunci PTO se transformă în CSO, iar VIO în formă personală.

LDVE

- 1) Dacă SVR precede SNT, atunci substantivul NOO sau grupa lui stau în acuzativ;
- 2) FNO din propoziția principală determină și numărul VNO din propoziția subordonată.

La nivelul segmentelor Nominal+Tematice SNT:

- 1) Genul, numărul, modul de articulare (hotărât, nehotărât) și cazul pentru toți membrii flexibili ai segmentului Nominal+Tematic (SNT) se determină după substantiv (NOO);
- 2) Determinarea cazului:
 - a) dacă SNT precede SVR, atunci toate părțile de vorbire stau în nominativ;
 - b) dacă SNT urmează după SVR, atunci toate părțile de vorbire stau în acuzativ.
- 3) Genul tuturor părților de vorbire flexibile se determină de substantivul (NOO) SNT;
- 4) Determinarea modului de articulare (hotărât/nehotărât):
 - a) NOO a SNT devine hotărât:
 - dacă în varianta engleză SNT a fost introdus de AHO (articol hotărât), funcția căruia a fost redată de NOO prin articularea definită a substantivului;
 - dacă este urmat de FSO (pronumele posesiv). Celelalte părți flexibile ale SNT (EQO, ECO, FIO) se acordă cu NOO.
 - b) NOO a SNT devine nehotărât:
 - dacă este precedat de ANO (articol nehotărât);
 - dacă sunt prezenți în SNT factori indicați mai sus (vezi reg. 4 a. Punctele I și II).
- 5) Echivalentul de traducere este determinat prin filtrarea tuturor variantelor propuse de dicționarele specializate și nespecializate (corespondența de afaceri în limbile română și engleză) care sunt supuse analizei corespunzătoare;
- 6) Varianta de traducere NOO și celelalte părți de vorbire flexibile a SNT este determinată de indicele genului, numărului, modului de articulare (hotărât/nehotărât) și cazului.
- 7) Cazul ours: dat fiind SNT (*ours/yours*) este reprezentat prin APO

și FSO, adică fără prezența NOO, devine imposibilă determinarea automată a genului, articulării și numărului, cazul este determinat dreptul că menționatul SNT precede SRT, adică stă în nominativ. Pentru a înlătura situația creată este necesar detectarea NOO care ar putea defintiva genul, articularea și numărul. El se găsește în SRV care urmează SNT (*ours*): (*ours is a very competitive market*). Se realizează operația dislocării NOO (*market*) în SNT: *our market* și rearanjamentul SVR: *is very competitive*. Articolul nehotărât (ANO) nu mai poate fi utilizat deoarece prin îmbinarea FSO cu NOO exclude prezența ANO în SNT.

La nivelul segmentelor Prepozițional-Nominal-Determinative SPND:

- 1) SPND este introdus de PSO care determină cazul NOO ale segmentului. Dat fiind diferitele interpretări posibile ale PSO depind de regulă de NOO din segmentul precedent (SNT, APND) este necesar de a prelucra informația: care variantă de traducere cere (impune) NOO precedent și ca urmare și cazul;
- 2) Partea a doua (după PSO) a SPND reprezintă partea nominală, modul căreia corespunde SNT (sunt cazuri adăugătoare referitor la ordinea de cuvinte);
- 3) Dacă NOO este urmat de ZOO care îl definitivează, atunci el este articulat definit (hotărât);
- 4) Dacă NOO este precedat de ZOO, care nu-l definitivează, atunci el este articulat nedeterminat (nehotărât);
- 5) Articolul posesiv sau genitival (APO) apare la pronumele posesiv (*your* – FSO) și la substantivul în genitiv (NOO – *samples*). El se acordă în gen și număr cu substantivul regent (al, a, ai, ale, alor). Prezența în segmentul precedent (SNT) a îmbinării NOO + EQO (*bună calitate*) cere păstrarea APO cu traducerea *alor* (tuturor mostrelor DV);
- 6) Dacă NOO este precedat de FIO, atunci articularea sa depinde de sensul FIO dat fiind FIO hotărât, altfel nehotărât:
 - a) dacă NOO este precedat de *all, either (neither)*, atunci NOO devine hotărât (determinat);
 - b) dacă NOO este precedat de celălalt FIO el devine nehotărât (nedeterminat).
- 7) Dacă MOO (numeralul cardinal) precede NOO, atunci el rămâne pe

- locul său fără schimbare;
- 8) Traducerea PSO *at* în îmbinarea cu NOO (prețuri) este determinată ca la (*at prices – la prețuri*);
 - 9) Dacă SPND introdus de PSO (*from*) este urmat de alt SPND (*to*), ambele PSO fiind urmate de ZOO, atunci în traducere ele sunt îmbinate: între 7 și 10%;
 - 10) Dacă se detectează îmbinarea PSO + NOO (*at prices*) care nu este precedată sau urmată de determinant (hotărât/nehotărât), atunci se impune articularea nedeterminată, adică nehotărâtă;
 - 11) Îmbinarea *at lest* (RAO) se traduce *cel puțin*;
 - 12) PSO *with* ce acuzativ și articularea definită a NOO;
 - 13) Dacă două substantive urmează unul după altul, atunci primul devine adjectiv și se schimbă cu locul (NOO – NEO), traducerea este fixată *de prețuri*;
 - 14) Dacă PSO (*to*) este precedat în segmentul precedent (SVR) (*look forward*) care cere în engleză prezența prepoziția *to*, atunci în limba română PSO este omis, iar NOO stă în acuzativ;
 - 15) Ordinea de cuvinte: dacă NOO este însoțit în continuarea segmentului de FSO și EQO, atunci primul FSO precede EQO;
 - 16) Lipsa unui determinant impune folosirea articulării indefinite pentru NOO (*vedere*).

Operațiunile automate de realizare a seturilor de reguli corespunzătoare cerințelor LDVE și LDVI permit fixarea totalului de caracteristici morfologic-sintactice în ordinea lor de succesiune pentru fiecare parte de vorbire în română și utilizarea lor la alegerea formei variantei echivalentului român de traducere.

De exemplu, pentru substantiv aceste caracteristici sunt: partea de vorbire (substantiv – sub.), genul (masculin – m, feminin – f, neutru – n), numărul (singular – sg, plural – pl), cazul (nominativ – nt, genitiv – gt, dativ – dt, acuzativ – at).

Pentru verb procesul de fixare finală a caracteristicilor îndeosebi constante este mai complicat, cu toate că și el se realizează în limitele propoziției analizate și segmentelor detectate folosindu-se de LDVE și LDVI.

Pentru fiecare parte de vorbire flexibilă se elaborează modelul paradigmelor de declinare și conjugare care permit computerului să compare datele fixate în modelul paradigmatic menționat mai sus. Coin-

cidența în comparație va permite determinarea automată a variantei de traducere (W1, W2, W3...).

De exemplu, pentru substantivul *ofertă – proposal* sunt fixate caracteristicile: sub. f. sg. ht. nt. (at). Comparându-le cu cele fixate în paradigma de declinare a acestui substantiv:

W1	sub	F	sg	ht	nt (at)	oferta
W2	sub	F	sg	ht	gt (dt)	(a) ofertei
W3	sub	F	sg	ht	nt (at)	(o) ofertă
W4	sub	F	sg	ht	gt (dt)	(a) (unei) oferte

Computerul alege W1, adică *oferta*.

Pentru adjectivul *bună – good* din segmentul *calitatea bună (the good quality)* sunt fixate caracteristicile: adj. atr. f. sg. nt. at. (nt). Realizând comparația cu caracteristicile fixate în paradigma sa de declinare:

W1	adj. atr.	m/n	sg	Nt	nt (at, gt, dt)	bun
W2	adj. atr.	f	sg	Nt	nt (at)	bună
W3	adj. atr.	f	sg	Nt	gt (dt)	bune
W4	adj. atr.	m	pl	Nt	nt (at, gt, dt)	buni
W5	adj. atr.	f/n	pl	Nt	nt (at, gt, dt)	bune

Computerul alege W2, adică *bună – calitate bună*.

Realizarea consecutivă a operațiunilor de traducere automată propusă de noi a scrisorii de afaceri engleze în limba română fără post redactor este reflectată în tabelul nr. 4.

Traducerea și schimbările intervenite în sistemul de segmentare și schematizare a propozițiilor engleze după traducerea lor în română

Tabelul nr. 4

Nr. seg	Tipurile de segmente				
	0. semnele de punctuație și conjuncțiile – CPS	1. Verbal-rematic – SVR	2. Nominal-tematic – SNT	3. Prepozițional-nominal-determinativ – SPND	4. Adverbial-determinative – SAD
1		Am primit în timp			
2			oferta Dvs.		
3				pentru livrarea	

4				de textile	
5				împreună cu mostrele și lis- tele de prețuri	
6					
1	Deși				
2		apreciem			
3			calitatea bună		
4				a tuturor mostrelor dvs.	
5					
6		considerăm			
7			prețurile		
8				la eșantioanele 235 și 236	cam ridicate
9					
10				pentru piața noastră.	
1		Trebuie totodată să subliniem			
2	că				
3			materiale foarte bune		
4		sunt acum accesibile			
5				din alte surse	
6				la prețuri	
7				între 7	
8				și 10 %	
9					mai mici de- cât ale Dvs.
1	Întrucât				
2		avem			
3				În vedere	
4			O comandă		
5				de aproximativ 500 yarzi	
6				de fiecare mostră	

7					
8		am dori să știm			
9	Dacă				
10		Ne-ați putea oferi			
11			un preț ce- va mai mic		
12					cel puțin
13				pentru cele două eșantioane sus menționate	
14					
15	ținând seama de faptul că				
16			Piața noastră		
17		este foarte competitivă			
18	și				
19			un preț avantajos		
20		este esențial			
1		sperăm			
2	Că				
3		veți fi dispuși să satisfaceți			
4			Condițiile noastre		
5	și				
6		Așteptăm cu interes			
7				răspunsul dvs. prompt.	

Schema succesiunii operațiunilor de traducere automată a corespondenței de afaceri engleze în limba română fără post-redactor
Textul englez

1. Detectarea și fixarea părții de vorbire sau semnelor de punctuație

- și atribuirea codului pe baza datelor de codificare.
2. Segmentarea automată a propoziției pe baza tabelelor de indicii.
 3. Schematizarea segmentelor fixate în propoziție.
 4. Dezambiguarea părților de vorbire segmentate.

Textul corespunzător român

5. Decodificarea (dacă ea este necesară).
6. Stabilirea ordinii de cuvinte.
7. Filtrarea și fixarea echivalentului de traducere.
8. Dezambiguarea părților de vorbire segmentate.
9. Determinarea legăturilor determinante de valența exterioară.
10. Determinarea legăturilor determinante de valență interioară.
11. Fixarea finală a caracteristicilor morfologo-sintactice a părților de vorbire segmentare.
12. Alegerea automată a variantei de traducere.
13. Fixarea rezultatului – traducerea textului englez în limba română fără post-redactor.

Scrisoarea de afacere engleză supusă analizei

We have duly received your proposal for the Supply of textiles with pattern-book and price lists.

While appreciating the good quality of all your samples, we find the price of patterns 235 and 236 rather high for our market.

We have also to point out that very good similar materials are now available from other sources at prices from 7 to 10% below yours.

As we have in mind an order of some 5000 yards for each sample, we would like to know if you could give us a slightly better price at least for the two above-mentioned patterns, considering that ours is a highly competitive market and a keen price is essential.

We trust that you will be willing to meet our terms and look forward to your prompt reply.

Observații generale:

1. Experimentul propus de traducere automată a scrisorilor de corespondență de afaceri engleză în limba română este dedicat simulării procesului de analiză și sinteză morfologico-sintactico-semantică realizat în creierul uman la generarea și perceperea acestui text specializat și formatizat în corespondere cu cerințele computerului prin algoritimizare și programare.

2. Luând în considerație arhicomplexitatea acestui proces el a fost divizat și subdivizat în unități și subunități de ordin lingvistic care permit rezolvarea succesivă a problemei, când soluția precedentă favorizează cele ulterioare.
3. Este una din primele lucrări de acest gen care se adresează direct la funcționarea mecanismului sinergetic lingvistic. Ca rezultat am desemnat existentele relații interioare de interdependență dintre fenomenele lingvistice și psihice care stau la baza funcționării acestui mecanism, în timp ce ideile, gândurile noastre, sunt produse într-o formă ideală și sunt proprietăți a materiei organizate într-un mod deosebit, raționamentele, noțiunile, concluziile și însăși limba naturală sau cea artificială sunt produse într-o formă concretă și materializată – structuri gramaticale și fondul vocabular.

Interdependența între limbă ca fenomen spiritual, pe de o parte, și materializat, pe de altă parte, ne-a permis pe bază procesării rezolvarea problemei traducerii automate a textului specializat fără post redactor.

BIBLIOGRAFIE

1. Chiriacescu A., Mureșan L., Barghiei V. *Corespondența de afaceri în limbile română și engleză*. București: Teora, 2000.
2. Cijacovschi V. *Aspectele semantice și comunicative ale procesării automate a textelor specializate*. Teză de doctor habilitat în științe filologice. St. Peterburg, Universitatea de Stat, 1988.
3. Cijacovschi V., Popescu A. *Particularitățile statistice și structurale ale scrisorilor de afaceri engleze*. Chișinău: ULIM, 1998.
4. Cijacovschi V., Popescu A. *Manifestarea mecanismului sinergetic la realizarea procesării automate a textului tehnico-științific*. Modele științifico-filologice. Chișinău: ULIM, 2003.
5. Cijacovschi V., Paladi A., Popescu A. *Modelarea informativ-lingvistică – instrument de cercetare a funcționării mecanismului sinergetic verbal*. In: Simposia Professorum. Chișinău: ULIM. 2002, pp. 135-140.
6. Piotrovsky R. *Automatul lingvistic și fundamentarea sa verbal-cognitivă*. Minsk, 1999.