

UNIVERSITATEA „ALEXANDRU IOAN CUZA” DIN IAȘI

FACULTATEA DE LITERE

ȘCOALA DOCTORALĂ DE STUDII FILOLOGICE

Probleme de stil în traducerea automată a textului literar

(Un studiu român-francez cu posibilități de extindere)

Rezumatul tezei de doctorat

Conducători de doctorat:

Prof. univ. dr. habil. Magda Jeanrenaud

Prof. em. dr. Dan Cristea, membru corespondent al Academiei Române

Doctorandă:

Lazăr N. Alexandra (căs. Ilie)

UNIVERSITATEA "ALEXANDRU IOAN CUZA" DIN IAȘI

Școala Doctorală de Studii Filologice

ANUNȚ

La data de 26.05.2025, ora 12, în sala „Vespasian Pella”, Casa Universitarilor doamna LAZĂR (căs. ILIE) N. ALEXANDRA va susține, în ședință publică, teza de doctorat cu titlul *„Probleme de stil în traducerea automată a textului literar. (Un studiu român-francez cu posibilități de extindere)”*, în vederea obținerii titlului științific de doctor în domeniul FILOLOGIE

Comisia de doctorat are următoarea componență:

Președinte

prof. univ. dr. habil. Antonio PATRAȘ, Universitatea “Alexandru Ioan Cuza” din Iași

Conducători de doctorat

Prof. univ. dr. habil. MAGDA JEANRENAUD, Universitatea „Alexandru Ioan Cuza” din Iași

Prof. univ. em. dr. habil. DAN CRISTEA, Universitatea „Alexandru Ioan Cuza” din Iași

Referenți

C.S. I, dr. LUMINIȚA BOTOȘINEANU, Institutul de Filologie „Al.Philippide” Iași, Filiala Academiei Române

C.S. II dr. Verginica MITITELU, Institutul de Inteligență Artificială „M. Drăgănescu”, Academia Română, București

prof. univ. dr. habil. Ioan MILICĂ, Universitatea „Alexandru Ioan Cuza” din Iași

Vă invităm să participați la ședința de susținere a tezei.
Teza poate fi consultată la Biblioteca Facultății de Litere

Data afișării: 06.05.2025

CUPRINS

Introducere	1
Problematica.....	1
Obiective	2
Metodologie	2
Capitolul I: Cadrul teoretic	9
1.1. Studii de lingvistică generală.....	10
1.2. Studii de teoria literaturii	37
1.3. Studii de traductologie	43
1.4. Prelucrarea limbajului natural.....	74
1.5. Traducerea automată	102
Capitolul II: Antrenarea unui sistem de traducere folosind <i>Moses</i>	115
2.1 Crearea corpurilor de antrenament	115
2.3. Antrenarea modelelor de traducere automată.....	116
2.4. Crearea memoriilor de traducere.....	119
2.5. Crearea modelelor de limbă	119
2.6. Traducerea automată și evaluarea traducerii	120
Capitolul III: Analiza textelor literare.....	123
3.1. Analiza contextului extrinsec	130
3.2. Analiza contextului intrinsec.....	134
3.2.1. Antroponimele – Personajele.....	134
3.2.2. Indicii spațiale exterioare – Toponimele și odonimele	141
3.2.3. Indicii spațiale interioare	143
3.2.4. Indiciile temporale.....	144
3.3. Analiza și crearea modelelor actanțiale.....	145
3.3.1. Analiza manuală:	145
3.3.2. Analiza automată	149
Capitolul IV: Traducerea și evaluarea	155
4.1. Studiu de caz 1: <i>O noapte furtunoasă</i> de Ion Luca Caragiale.....	158
4.1.1 Analiza cantitativă a traducerilor	158
4.1.2. Analiza stilistică a traducerilor	169
Nivelul fonetic	169

Nivelul lexical	187
Nivelul morfo-sintactic	222
Nivelul stilistic	241
4.2. Studiu de caz 2: <i>Ion...</i> de Ion Luca Caragiale	253
Nivelul fonetic	253
Nivelul lexical.....	255
Nivelul morfo-sintactic.....	259
Nivelul stilistic.....	260
4.3. Studiu de caz 3: <i>En passant</i> de Raymond Queneau.....	262
4.3.1. Textul sursă.....	263
4.3.2. Traducerea umană	272
4.3.3. Traducerea automată	277
4.3.3.1. Analiza cantitativă a traducerilor.....	277
4.3.3.2. Analiza stilistică a traducerilor	286
Nivelul fonetic	286
Nivelul lexical.....	288
Nivelul morfo-sintactic.....	290
Nivelul stilistic.....	293
Capitolul V: Concluzii finale.....	301
BIBLIOGRAFIE.....	307

Lista abrevierilor

A = Adjuvant
AP = propoziție apozitivă
AT = propoziție atributivă
BLEU = Bilingual Evaluation Understudy
BP = Brevity Penalty
CAT = Computer-assisted translation (traducere asistată de calculator)
CD = propoziție completivă directă
CM = propoziție completivă de mod
CT = propoziție completivă de timp
D1 = Destinator
D2 = Destinatar
DL = DeepL
GT = Google Translation
LS = limba sursă
LSTM = Long Short-Term Memory
LȚ = limba țintă
MERT = Minimum Error Tate Training
MMT = Multi-lingual Machine Translation (Traducerea Automată Multimodală)
MT = Microsoft Translator
NLP = Natural Language Processing (Procesare Limbajului Natural)
NMT = Neural Machine Translation (Traducerea Automată Neurală)
O = Obiect
OCR = Optical Character Recognition (Recunoaștere Optică a Caracterelor)
Op = Opozant
P = propoziție
PP = propoziție principală
RNN = Recurrent Neural Network (Rețele Neuronale Recurente)
S = Subiect
SMT = Statistical Machine Translation (Traducerea Automată Statistică)
t.n. = traducerea noastră
TA = traducere automată
TS = text sursă
TȚ = text țintă
TU = traducere umană
TU¹ = traducerea umană realizată de Simone Roland și Valentin Lipatti
TU² = traducerea umană realizată de Eugen Ionescu și Monica Lovinescu
TU³ = traducerea umană realizată de Paola Bentz-Fauci
XML = Extensible Markup Language

Index de tabele

Tabel 1. Probleme de traducere.....	58
Tabel 2. Clasificarea tipurilor de traducere la C. Nord	70
Tabel 3. Corpusuri paralele OPUS Corpus 2019 română-franceză	83
Tabel 4. Clasificarea corpusurilor după criteriul tipului de text	85
Tabel 5. Corpusuri paralele OPUS Corpus 2024 română-franceză	88
Tabel 6. Cuvinte frecvente în traduceri umane pentru O noapte furtunoasă.....	162
Tabel 7. Cuvinte frecvente în traduceri automate pentru O noapte furtunoasă.....	165
Tabel 8. Ocurențele numelor personajelor în textul sursă En passant și numărul de replici	264

Index de figuri

Figura 1. Modelul lingvistic la Eugeniu Coșeriu	12
Figură 2. Modelul comunicării la K. Bühler.....	14
Figură 3. Modelul comunicării la C. Shannon.....	15
Figură 4. Modelul comunicării la R. Jakobson	16
Figură 5. Model de traducere	39
Figură 6. Procedee de traducere.....	50
Figură 7. Procesul de analiză textuală.....	81
Figură 8. Număr de tokenuri disponibile în OpusCorpus pentru limba franceză (2019)	89
Figură 9. Număr de tokenuri disponibile în OpusCorpus pentru limba franceză (2024)	89
Figură 10. Număr de tokenuri disponibile în OpusCorpus pentru limba română (2019)	90
Figură 11. Număr de tokenuri disponibile în OpusCorpus pentru limba română (2024)	90
Figură 12. Tipuri de texte disponibile în OpusCorpus pentru limba română (2024).....	93
Figură 13. Tipuri de texte disponibile în OpusCorpus pentru limba franceză (2024)	93
Figură 14. Interfață Romtext 2019.....	94
Figură 15. Frantext 2019.....	96
Figură 16. Frantext 2024.....	97
Figură 17. Reprezentarea modelului actanțial propusă de A. Ubersfeld.....	125
Figură 18. Reconstruirea traseului urmat de Jupân Dumitrache în seara de „lăsata secului”	142
Figură 19. Model actanțial 1: Jupân Dumitrache - subiect	146
Figură 20. Model actanțial 2a: Chiriac - subiect.....	147
Figură 21. Model actanțial 2b: Chiriac - subiect.....	147
Figură 22. Model actanțial generat folosind Python graphviz.....	149
Figură 23. Relația între personaje folosind Python graphviz.....	150
Figură 24. Temele centrale generate prin Python graphviz	150
Figură 25. Scene centrale generate prin Python graphviz.....	151
Figură 26. Personaje și teme centrale generate cu Python graphviz.....	151
Figură 27. Relațiile între personaje generate cu Python graphviz	152
Figură 28. Relațiile între personaje generate cu Python graphviz	152
Figură 29. Teme centrale generate cu Python graphviz	153
Figură 30. Teme centrale generate cu Python graphviz	153
Figură 31. Scene centrale generate cu Python graphviz	153
Figură 32. Scene centrale generate cu Python graphviz	154
Figură 33. Frecvența de utilizare a termenului „jupân” în textul sursă	190
Figură 34. Frecvența de utilizare a termenului „jupân” în traducerile automate	191
Figură 35. Dinamica menționării personajelor în TS.....	278
Figură 36. Dinamica menționării personajelor în traducerea Google Translate	280
Figură 37. Dinamica menționării personajelor în traducerea Microsoft Translator.....	281
Figură 38. Dinamica menționării personajelor în traducerea DeepL.....	283
Figură 39. Dinamica menționării personajelor în traducerea Chat GPT	284
Figură 40. Dinamica menționării personajelor în traducerea umană.....	285

Introducere

Problematica

Cercetarea de față s-a conturat în jurul ideii de colaborare interdisciplinară. Scopul nostru a fost acela de a identifica, de a clasifica greșelile care apar în traduceri literare realizate prin utilizarea sistemelor automate de traducere și de a găsi motivele pentru care traduceri automate actuale nu ating calitatea traducerilor umane. Am inclus și câteva recomandări pentru dezvoltatorii sistemelor de traducere automată în vederea îmbunătățirii calității traducerilor de texte literare. Studiul se adresează cercetătorilor preocupați atât de prelucrarea și de analiza textelor literare folosind instrumente digitale, cât și de elaborarea sistemelor informatice pentru prelucrarea textelor literare. De asemenea, studiul nostru este un instrument util și pentru cei care sunt în căutarea unei metode de evaluare orientată spre o stilistică a traducerilor literare realizate atât de traducători umani, cât și de sistemele de traducere automată.

Domeniul traducerii este, prin excelență, un domeniu al colaborării, al interacțiunii, al confruntărilor și al rezolvărilor, al soluțiilor găsite chiar și în situații de aparentă imposibilitate. Așa cum traducătorul uman reușește să depășească obstacole dificile întâlnite în procesul traducerii prin împletirea mai multor strategii, considerăm necesară o abordare pluridisciplinară în dezvoltarea sistemelor de traducere automată. Fără a fi lipsit de dificultate, demersul nostru nu este, totuși, imposibil. De altfel, traducerea însăși ne demonstrează acest lucru. Pe cât de anevoioasă a fost întreprinderea unui studiu care să pună în contact discipline independente care observă textul literar din perspective diferite și îl analizează prin mijloace proprii, pe atât de plină de satisfacție a fost activitatea propriu-zisă de cercetare. În lucrarea de față propunem un model de evaluare a traducerilor textelor literare bazându-ne pe teorii elaborate în lingvistica generală și aplicată, teoria literaturii, traductologie, lingvistica computațională, prelucrarea limbajului natural, traducerea automată. Modelul nostru de evaluare a traducerilor literare pornește de la constatările făcute în urma analizelor comparative realizate în trei etape. Mai întâi, am analizat comparativ textele originale și traduceri lor realizate de traducători umani. A urmat o analiză a textelor traduse prin sisteme informatice de traducere (*Google Translate*, *Microsoft Translator*, *DeepL* și *Chat GPT*) și textele originale, studiu completat cu o analiză comparativă între traduceri

automate și traduceri umane. Precizăm că aceste analize s-au efectuat asupra nivelului fonetic, lexical, morfo-sintactic și stilistic, după cum propun și lucrările de stilistică. Textul sursă a fost analizat și din perspectiva celor patru niveluri menționate.

Obiective

Scopul nostru a fost acela de a analiza erorile apărute în traduceri automate, de a le clasifica după criteriul nivelului de limbă și de a concepe un model de evaluare al traducerilor automate ținând cont de teoriile elaborate în studiile de traductologie (tipul de text, scopul traducerii, publicul țintă etc.). Dacă traducătorului îi revine sarcina de a descifra enigmele pe care autorul le-a meșteșugit și inserat dinadins în text, de a identifica stilul autorului, de a reconstrui textul cu instrumentele limbii sale, de a făuri și a le încifra cu aceleași chei pe care autorul le-a utilizat în realizarea textului său, evaluatorului îi revine analiza modului în care traducătorul a reușit să-și îndeplinească sarcina. Știind că stilul și creativitatea unui traducător pot altera stilul autorului textului sursă și că aceasta poate influența calitatea unei traduceri, ne întrebăm cum poate fi măsurată calitatea unei traduceri automate? Dar stilul sau creativitatea traducătorului sau al sistemului automat de traducere? Există un stil al traducătorului? Cum să „învățăm” un sistem automat să identifice și să recreeze stilul? Cum ar putea fi acesta „creativ”? Există creativitate? Ce este, de fapt, stilul și creativitatea? Pot fi transpuse, traduse și măsurate acestea într-un mod obiectiv? Dar automat? Sunt câteva dintre întrebările la care ne-am propus să răspundem în acest studiu sprijinindu-ne pe teoriile ce au contribuit la dezvoltarea traducerii automate, identificând, totodată, și modul în care acestea au contribuit la starea actuală a traducerii automate a textului literar.

Metodologie

Realizarea acestui studiu s-a sprijinit pe patru piloni importanți, traductologie, lingvistică generală, critică literară și lingvistică computațională, pe care-i considerăm definitorii pentru analiza traducerilor automate. Identificarea și analiza problemelor de stil în traduceri automate

pretinde cercetătorului să urmărească, de asemenea, și metodologia de cercetare în domeniul stilisticii, motiv pentru care am completat abordarea traductologică cu cea stilistică. Evaluarea traducerilor literare și elaborarea unui model de evaluare a traducerilor automate ar fi fost imposibile fără investigarea studiilor de profunzime care stau la baza creării și evaluării textului literar. Desprinsă de alte discipline conexe, de mai bine de 50 de ani, traductologia a devenit o disciplină de studiu cu statut propriu și individual, sub umbrela mai largă a lingvisticii generale. În paralel, traducerea automată este o ramură de studiu desprinsă din studiile de lingvistică computațională și de prelucrare a limbajului natural. Așa cum traductologia nu se poate dezlipi întru totul de lingvistica generală, progresul traducerii automate este direct influențat de evoluția și rafinarea sistemelor de prelucrare a limbajului natural. Dacă lingvistica generală ne oferă direcțiile necesare pentru studiul formei, studiile de teorie a literaturii ne-au oferit baza pentru investigarea conținuturilor. Prin lucrarea de față propunem o combinație a studiilor de traductologie cu studiile de traducere automată cu scopul de a îmbunătăți calitatea traducerilor automate actuale ale textelor literare pentru perechea de limbi română-franceză.

Traductologia a stat la baza realizării studiului, delimitându-l de studiile de lingvistică sau de teorie a literaturii, fără însă a-l îndepărta de acestea. În traducerile analizate am identificat tehnicile de traducere, defectele și calitățile traducerilor despre care am discutat în secțiunea teoretică a cercetării noastre. Ne-am propus să observăm care dintre aceste tehnici sunt folosite și în practica traducerilor automate și ce anume din teoriile de traducere dezvoltate de-a lungul timpului a reușit să pătrundă în aria prelucrării limbajului natural folosind instrumente digitale.

În ceea ce privește studiile de lingvistică generală, am prezentat cadrul teoretic în care s-au dezvoltat metodele de cercetare a limbii, printre care menționăm analiza funcțională, analiza distribuțională, analiza în constituenți imediați, analiza în trăsături distinctive. Aceste metode analitice stau la baza interpretării limbajului natural prin sisteme digitale. Studiile de stilistică au reprezentat baza analizelor pentru textele sursă și un model de lucru. Am argumentat apariția în textul sursă a fenomenelor de natură fonetică, lexicală, morfo-sintactică și stilistică prin instrumente specifice fiecărui nivel lingvistic în parte. Explicarea selecțiilor autorului ne-a ajutat să identificăm dacă și cum au fost acestea recuperate, transformate, transmise în traduceri.

Studiile de teorie a literaturii au fost indispensabile pentru scopul acestui studiu. Ne-am propus o evaluare a traducerilor literare, obiectiv realizabil printr-o analiză critică a textului, prin

metoda comparativă bilingvă. Într-o primă etapă a evaluării am analizat critic textele literare, pentru fiecare nivel lingvistic. Am extras punctele de maximă intensitate stilistică a căror transformare am observat-o în traduceri. De aceea, a doua etapă a evaluării a constatat în confruntarea textului sursă cu traducerile sale umane și automate. În ce-a de-a treia etapă a evaluării am comparat traducerile între ele având ca text suport de verificare textul sursă și am extras greșelile de traducere prezente în traducerile automate, prin raportare la traducerile umane și la textul sursă.

În ceea ce privește studiile privind prelucrarea limbajului natural am stabilit cadrul teoretic și istoric al progreselor înregistrate până la apariția și dezvoltarea traducerii automate. Am identificat etapele pe care le parcurge un text, de la textul imprimat, în format fizic, până la traducerea lui automată, în format digital, prin sisteme automate de traducere. Am discutat despre importanța corpusurilor, a memoriilor de traducere și a intervenției specialiștilor umani în dezvoltarea acestor sisteme. Ne-am referit la situația actuală pentru perechea de limbi română-franceză.

După această introducere generală, pătrundem în laboratorul lucrării de față. Titlul, *Probleme de stil în traducerea automată a textelor literare. (Un studiu român-francez cu posibilități de extindere)*, vorbește de la sine. Aveți în față un studiu stilistic pentru perechile de limbi română-franceză în care urmărim tipurile de erori apărute în traducerea automată a textelor literare, clasificate după nivelurile limbii. Am extras din lumea textului literar un tip special, pe cel al textelor dramatice și al textelor narative bogate în dialoguri. Motivația s-a sprijinit pe structura tehnică a dialogului, bazat, în linii generale, pe sistemul întrebare-răspuns și pe schimbul alternativ de replici. Cu alte cuvinte, nu am cercetat limba textelor literare, ci limbajul înscris în universul lor. Printre preocupările noastre s-a numărat și identificarea diferențelor dintre actul comunicării verbale prezent în vorbire și actele discursive prezente în limba scrisă a textelor literare sub formă de dialoguri. Narațiunea și descrierea nu au fost neglijate, fiind analizate ca elemente cu rol de întregire a imaginii de ansamblu a operei literare. Am ales doi maeștri ai limbajului preocupați de mecanica și funcționarea limbii, unul din cultura românească, unul din cea franceză: Ion Luca Caragiale și Raymond Queneau. Ne-au interesat texte traduse în limba franceză, respectiv română, texte care au constituit corpusul nostru de investigație. Ar putea avea ceva în comun Caragiale și Queneau, un clasic și un modern? Ce ne-a făcut să-i aducem împreună pe acești doi mari scriitori

a fost interesul lor pentru intimitatea profundă a limbii și pentru crearea mecanismelor proprii de exprimare a personajelor imaginate și nu în ultimul rând, o predilecție a cercetătorului pentru aceștia.

Tiparele de construcție a ansamblului textual, al personajelor, al modului lor de exprimare au reprezentat un factor decisiv în alegerea noastră. Textele pieselor de teatru scrise de I. L. Caragiale cuprind elemente tehnice care reglează și dozează conținutul, textele lui Raymond Queneau, deși abstracte, au la bază tehnici de realizare matematice, după cum aflăm din programul *OuLiPo*, al cărui fondator a fost alături de François Le Lionnais. Similitudinile de ordin metatextual pe care le-am identificat în textele analizate au reprezentat, de asemenea, motive care au sprijinit alegerea noastră.

Deși aparțin unor epoci diferite, abilitatea de construire a limbajului, măiestria în exploatarea laturii tehnice a limbajului și a capacității lui performatoare sunt excepțional reprezentate în operele acestor doi mai autori. Alegerea noastră a fost determinată și de faptul că nu doar limba este explorată, ci limbajul în ansamblul său. Urmărind haosul limbajului până în cele mai intime detalii, atât Caragiale, cât și Queneau au reușit cu măiestrie să-l redea în forma coerentă a textului literar. Limba vie pe care Caragiale a așternut-o în paginile sale face posibilă observarea unui sistem diferit de cel al limbii literare standard. Predilecția lui a fost pentru limbajul cules „de pe ulițele murdare ale mahalalelor” Bucureștilor (Vitner, 1953, p. 46). Autorii caută nu frumusețea și perfecțiunea exprimării, ci limbajul inestetic, imperfect, uneori chiar brutal sau grotesc redat într-o estetică a urâtului. Queneau face din limbaj un obiect de studiu, cu care negociază la nesfârșit. Roland Barthes afirma despre literatura lui Queneau că „este o categorie a vorbirii, deci a existenței și care privește întreaga umanitate” (Barthes, 1964, p. 134).

Mai mult, legătura mai puțin investigată dintre Caragiale și Queneau a făcut-o și scriitorul suprarealist Eugen Ionesco, de altfel, membru al *OuLiPo* care, împreună cu Monica Lovinescu, aduce în „lumina reflectoarelor” limbii franceze trei piese ale lui Caragiale: *O noapte furtunoasă*, *Conu' Leonida față cu reacțiunea* și *O scrisoare pierdută* (Caragiale, 1994). În plus, am constatat că un contact anterior între cei doi autori (desigur, dintre Queneau spre Caragiale) trebuie să fi existat, de vreme ce Queneau vorbește, folosind termenii lui Caragiale în romanul *Pierrot mon ami*, despre „mitocans”, „mocofans” și „badarans” (Queneau, 1945, p. 100), referindu-se la acei „gură-cască” ai mediului urban. Nu limbajul elegant, delicat, manierat, rafinat apare în textele celor

doi scriitori, ci un limbaj specific pentru personajele lor, căutat și meșteșugit special pentru ele, un limbaj ce exprimă natura lor umană. Paul Zarifopol spunea despre Caragiale că „s-a cheltuit în literatură vorbită (...) constrâns la muncă tăcută ce-l face să-și lucreze scrisul cu o trudă înverșunată (...), obsedat, și când tace, de efectele sentimentale și estetice ale vorbei rostite în gura mare” (Zarifopol, 1922, p. 395). Personajele lui Queneau, de altfel, folosesc același limbaj familiar, pe alocuri grosier, viu, într-o prefacere perpetuă. În cuprinsul textelor acestor doi autori am identificat punctul de congruență între vorbire și limbă, locul în care natura limitată, rigidă și matematică a limbii abstracte este impecabil pusă în serviciul vorbirii nelimitate, flexibile și expresive, locul actului concret de manifestare al tuturor invariantelor în variante încălcând regulile de transformare standard în trecerea structurilor de adâncime în structuri de suprafață.

Am abordat dialogul acestor texte literare și din perspectiva lingvistică și stilistică a vorbirii, nu doar a limbii care constrânge vorbirea. Ipoteza noastră este că un sistem automat de traducere ar putea gestiona mai bine un text omogen, format din fraze relativ scurte cu elemente care se repetă sistematic. În ceea ce privește textele „matematice” și jocurile tehnice de creare a literaturii propuse de Queneau, ne-am dorit să observăm în ce măsură a fost descifrată această intenție a autorului de către traducătorii umani, precum și identificarea elementelor care se regăsesc în traducerile în limba română.

Menționăm și faptul că studiul de față, fiind condiționat și mărginit de teoriile expuse în domeniile de studiu menționate, a suferit constant reglaje astfel încât să nu intre în contradicție cu niciuna dintre cele două mari zone la interferența căruia se află, lumea literelor și lumea informaticii.

În primul capitol am schițat cadrul general al cercetărilor anterioare a căror conlucrare a făcut posibilă investigarea traducerilor automate de texte literare. Ne-am concentrat atenția asupra aspectului practic, aplicat al acestor cercetări. Am inventariat studiile teoretice care au contribuit la starea actuală a ramurii de studiu a traducerii automate și am insistat asupra conceptelor operaționale care au avut un impact major în evoluția domeniului de cercetare. Am prezentat succint etapele prin care trece un text literar, de la formatul analog la cel digital, precum și o privire de ansamblu asupra proceselor necesare realizării unei traduceri automate: crearea de corpuri paralele, adnotarea, antrenarea modelelor de limbă și de traducere, evaluarea.

În cel de-al treilea capitol am realizat un experiment de antrenare a unui sistem de traducere folosind *Moses*. Am discutat despre importanța corpusurilor paralele și am creat un corpus de antrenament și modele de limbă, după care am antrenat două modele de traducere. La final am evaluat rezultatele obținute.

Al treilea capitol constă în selecția a două texte sursă din corpusul de analiză în vederea analizei lor, bazându-ne pe teoriile discutate în prima parte a studiului, în secțiunea dedicată studiilor de teorie a literaturii. Am observat atât nivelul macrostructural, cât și cel microstructural. Analiza microstructurală s-a orientat asupra nivelului fonetic, lexical, morfo-sintactic și stilistic. Pentru nivelul fonetic ne-a preocupat analiza punctuației și a onomatopeelor. În ceea ce privește nivelul lexical am acordat o atenție deosebită procedeelelor de îmbogățire a vocabularului, în special derivării cu sufixe augmentative și diminutive și împrumuturilor, numelor proprii, supranumelor, creațiilor proprii autorilor. Am extras în mod automat cele mai frecvente cuvinte și am discutat importanța lor în text. În privința analizei nivelului morfo-sintactic observăm valorile stilistice ale modurilor și ale timpurilor verbale, identificăm dezacordurile gramaticale și funcțiile lor în textele literare și, în cele din urmă, facem observații pe marginea topicii frazelor din textele sursă. Analiza nivelului stilistic constă în identificarea figurilor stilistice, a aluziilor folclorice, religioase, politice, culturale etc., a umorului, a ironiei. Completăm studiul cu observații privind tonul și registrul identificate în textele sursă selectate. Corelăm detalii aparținând contextului extrinsec cu elementele aparținând contextului intrinsec, particulare fiecărui text literar și fiecărui autor. În ceea ce privește contextul extrinsec discutăm câteva date despre autor și mediul socio-politic din care provine. Privind contextul intrinsec ne vor interesa antroponimele și indiciile spațiale și temporale. De asemenea, am inclus și o analiză pe baza modelelor actanțiale. Am făcut un experiment de realizare a modelelor actanțiale folosind sisteme digitale. Am observat modul în care sistemele automate reușesc să coreleze temele literare cu desfășurarea acțiunii, actanții și evoluția acestora.

Al patrulea capitol, cel mai amplu, constă în analiza traducerilor umane și automate. Propunem trei studii de caz plecând de la trei texte literare suport: *O noapte furtunoasă* și *Ion...* de Ion Luca Caragiale și singura piesă de teatru scrisă de Raymond Queneau, *En passant*. Primul text are trei versiuni de traducere umană, celelalte două texte nu. Pentru cel de-al treilea text realizăm o traducere în limba română pe care o vom confrunța cu traducerile automate. În toate cele trei

studii de caz observațiile noastre se vor concentra în jurul celor patru niveluri lingvistice menționate.

În fine, în ultima secțiune a studiului nostru am extras concluziile în urma tuturor experimentelor efectuate asupra evaluării traducerilor automate. Prin cercetarea noastră dorim să atragem atenția asupra importanței integrării studiilor filologice în procesul de elaborare a unui model de evaluare a traducerilor automate specializat în traducerea textelor literare, care să țină cont de strategiile de traducere, de conceptele lingvistice, de critica literară, de analiza discursului, de metodologia de evaluare a traducerilor făcute de traducători umani.

Cadrele teoretice prezentate în prima parte a studiului nostru ne-au permis să înțelegem importanța formei pe care o primește un text literar și funcțiile pe care le îndeplinește fiecare element constitutiv al comunicării artistice. Am demonstrat cu argumente lingvistice și de teorie literară că textul literar este o expresie situată în planul limbajului oral și că analiza lui necesită instrumente specifice, diferite de cele utilizate în analiza formală a limbii. De asemenea, am arătat că traducerea unui text literar impune o analiză critică prealabilă și am propus un model care ține cont atât de contextul extrinsec, cât și de cel intrinsec al operei. Am menționat importanța integrării acestor etape în sistemele de traducere automată.

Acest studiu poate fi dezvoltat prin crearea unui model de traducere specializat pe traducerea textelor literare care să integreze dicționare de termeni literari și de sinonime pentru perechea de limbi română-franceză. De asemenea, propunem un model de limbă antrenat pe texte literare. Antrenarea modelelor presupune adnotarea textelor la nivelul părților de vorbire și etichetarea entităților denumite pentru recunoașterea lor în context (NER). Acest lucru ar reduce cantitatea mare de greșeli gramaticale și ar rezolva greșelile generate în traducerea numelor de persoane, a denumirilor instituțiilor sau în reluarea substantivelor prin pronume. În cazul traducerilor automate de texte literare propunem crearea unui corpus paralel format din textele sursă și traducerile lor umane și crearea unor memorii de traducere. Însă toate aceste intervenții umane ar fi insuficiente fără o cantitate mare de texte literare, pe baza cărora să funcționeze un sistem de traducere automată.