

**Mihai-Bogdan Atanasiu • Anca-Diana Bibiri •
Emanuel Grosu • Alina Moroşanu •
Constantin Răchită**
(Editors)

CULTURAL DYNAMICS OF VALUES



EDITURA UNIVERSITĂŢII „ALEXANDRU IOAN CUZA” DIN IAŞI

2024

The book of proceedings of the conference
Interdisciplinary Perspectives in Humanities and Social Sciences
9th Edition: Rethinking Values in Interdisciplinary Research (27-28 October 2023)
Institute of Interdisciplinary Research
Department of Social Sciences and Humanities
"Alexandru Ioan Cuza" University of Iasi

Scientific Reviewers:

Prof. Adina Dornean, PhD habil.

Senior Researcher Marius Chelcu, PhD habil.

Language revision of texts in English:

Senior Researcher Anca-Diana Bibiri, PhD

Book editor:

Emanuel Grosu

Marius-Nicușor Grigore

Coverage: Manuela Oboroceanu

ISBN: 978-606-714-906-7

DOI: 10.47743/phss-2024

The authors are entirely responsible for the scientific contents of the texts herein published as well as for the fair use of the copyrighted works.

© Editura Universității „Alexandru Ioan Cuza” din Iași

700539 – Iași, Str. Munteni nr. 34, tel. (0232) 314947; editura@uaic.ro

www.editura.uaic.ro

Mihai-Bogdan Atanasiu is a senior researcher, director of the Department of Social Sciences and Humanities, Institute of Interdisciplinary Research, “Alexandru Ioan Cuza” University of Iași. He has a PhD in History, awarded in 2012 at the same university. His research activity focuses on the political, social and cultural history of Moldavia in the seventeenth and eighteenth centuries. Most of his scholarly contributions have focused on social history, genealogy, prosopography, history of the Church, history of mentalities, as well as on editing documentary sources.

Anca-Diana Bibiri is a senior researcher at the Department of Social Sciences and Humanities, Institute of Interdisciplinary Research, “Alexandru Ioan Cuza” University of Iași. She has a PhD in Philology, and a postdoctoral fellowship in linguistics from the same University, and her main areas of research are: prosody, phonetics and dialectology, computational linguistics, natural language processing, lexicography, and sociolinguistics. Co-editor of the PHSS Proceedings (2014-2019).

Emanuel Grosu is a senior researcher at the Department of Social Sciences and Humanities, Institute of Interdisciplinary Research, “Alexandru Ioan Cuza” University of Iași. He has a PhD in Philology. He published studies, exegesis and translations from medieval Latin authors (Paulus Diaconus, Dungalus Reclusus, Anselm of Canterbury, Marcus of Regensburg, Marco Polo), the diachronic evolution of central literary themes and motifs of medieval Latin culture constituting the main research direction. Co-editor of the PHSS Proceedings (2014-2019).

Alina Moroșanu is a senior researcher at the Department of Social Sciences and Humanities, Institute of Interdisciplinary Research, “Alexandru Ioan Cuza” University of Iași. She has a PhD in in Cybernetics and Economic Statistics, awarded in 2011 at the same university. Her research interest include: questionnaires development, healthcare management analytics, project management, statistical analysis, statistical software (R, SPSS), surveys, human resources analytics.

Constantin Răchită is a research assistant in the Department of Social Sciences and Humanities, Institute of Interdisciplinary Research, “Alexandru Ioan Cuza” University of Iași. He has a PhD in Philology. His primary research focuses on the translation and interpretation of ancient texts in Old Greek and Latin. Throughout his career, he has participated in various translation and editing projects. Currently, his research interests encompass interdisciplinary approaches to biblical and patristic texts, exploring issues related to translation, transmission, and their influence on contemporary society.

Descrierea CIP a Bibliotecii Naționale a României

Cultural Dynamics of Values / Mihai-Bogdan Atanasiu, Anca-Diana

Bibiri, Emanuel Grosu, - Iași : Editura Universității “Al. I. Cuza”, 2024

Conține bibliografie

ISBN 978-606-714-906-7

I. Atanasiu, Mihai-Bogdan

II. Bibiri, Anca-Diana

III. Grosu, Emanuel

Contents

Foreword (<i>Anca-Diana Bibiri</i>)	9
--	---

Plenary Conference

Human Enhancement: tehnologie versus teologie. Repere pentru o evaluare interdisciplinară a valorilor și posibilităților de devenire a umanului prin cunoaștere [Human Enhancement: Technology versus Theology. Landmarks for the Interdisciplinarity Evaluation of Human Values a Potential of Becoming Through Knowledge]

Pr. Andrei-Răzvan Ionescu	21
--	----

Philology

The Use of Artificial Intelligence (AI) in Linguistics. Case Study: Analysis of Linguistic Phenomena in the Novel *Ion* by Liviu Rebreanu

Cristina Bleorțu	35
-------------------------------	----

Traducerea automată a literaturii. O himeră încă vie? [Automatic Translation of Literature: A Still Living Chimera?]

Alexandra Ilie	51
-----------------------------	----

Kitsch. The Control and Faking of Aesthetic Value

Daniela Petroșel	77
-------------------------------	----

Valori perene în predarea romanisticii în spațiul universitar românesc [Perennial Values in Teaching Romance Studies in Romanian Universities]

Mihaela Secieru	93
------------------------------	----

Authentic vs. Pseudo Values

Paula-Andreea Onofrei	111
------------------------------------	-----

Medical Humanities Approached Through a Feminist Lens

Laura Ioana Leon	127
-------------------------------	-----

Explorări teoretice și suprapuneri terminologice. Romanul, obiect de reflecție și prim suport al teoriei genurilor [Theoretical Explorations and Terminological Overlaps. The Novel, as Object of Reflection and the First Support of the Genre Theory]

Alexandra Olteanu	141
--------------------------------	-----

Spectrele filiațiilor literare. Portrete ale generațiilor – Mircea Ivănescu și Radu Vancu [The Specters of Literary Filiations. Portraits of Generations – Mircea Ivănescu and Radu Vancu]	
Teodora Iurusiuc	165
Memoria comunismului în <i>Jurnalul unui jurnalist fără jurnal de Ion D. Sîrbu</i> [The Memory of Communism in Ion D. Sîrbu's <i>Journal of a Journalist without a Journal</i>]	
Oana-Elena Nechita	181
Language in the Church: Orthodox Religious Terminology in Polish and the Role of Translations in Establishing Lexical Norms	
Irina-Marinela Deftu	201
History & Theology	
<i>Non naturalibus desiderii, sed censibus aestimentur.</i> Piața romană de legume și fructe [<i>Non naturalibus desiderii, sed censibus aestimentur.</i> The Roman Vegetable and Fruit Market]	
Iulia Dumitrache	219
Un posibil clivaj între teoria asupra stilului și aplicarea acesteia în cazul Fericitului Augustin? Pluralitatea de stiluri în operele acestuia [A Possible Split Between the Theory of Style and Its Application in the Case of Augustine? The Plurality of Styles in Augustine's Works]	
Pr. Liviu Petcu	249
Conflicting Values during the French Wars of Religion (1562-1598): Loyalty to the King and Loyalty to God	
Andrei Constantin Sălăvăstru	261
Polemici teologice în <i>Praefatio paraenetica</i> a lui John Pearson (1613-1686) [Theological Polemics in <i>Praefatio paraenetica</i> by John Pearson (1613-1686)]	
Constantin Răchită	281
Moartea – o preocupare a vieții cotidiene în Iași veacului al XVIII-lea [Death – a Preoccupation of Everyday Life in the 18 th Century Iași]	
Mihai-Bogdan Atanasiu	301

Economics

The Use of Information and Communication Technologies in Business as a Value-Creating Tool: Analysis on the Case of SMEs in Romania Valentina Diana Rusu & Angela Roman	317
Evaluation of Hospital Financing in Romania: A Comparative Analysis pre- and post-Pandemic COVID-19 Mihai-Vasile Pruteanu & Alina Moroşanu	337
Green Jobs, Green Skills and Green Human Resource Management. An Analysis of Current Trends Silvia-Maria Carp & Ana-Maria Bercu	367
Is Security a Timeless Value? An Insight from International Relations Andreea-Cosmina Foca & Oana-Maria Cozma	381

Traducerea automată a literaturii. O himeră încă vie?

ALEXANDRA ILIE*

Abstract: *Artificial intelligence is about to reshape the world and redefine the relationship humans maintain with their professional activities. Recent machine translation systems, as well as the prior generation of statistical machine translation, rely on extensive aligned corpora and generate outputs whose quality can compete with human translations. This is particularly the case for technical and even medical texts. The role of the translator is to be questioned. Could algorithms potentially substitute this profession? Certain text types, such as legal and literary ones, remain resistant to machine translation. The main obstacle for these systems is their lack of direct, tangible interaction with the outer reality from which these texts emerge, hence the systems' inability to understand and interpret them. The ability of Transformer models is confined to analyzing intra-sentential context. In 2017, a new player entered the machine translation scene, DeepL, whose primary advantage lies in its integration of the Linguee database, an enormous parallel corpus of high-quality human translations. Generative natural language models include translation functionality. These models can be trained to identify or even replicate specific literary styles. Since literary works are replete with wordplay, extratextual references, and stylistic devices, we propose to investigate whether the task of decoding their message, particularly the author's intent (understood as the addresser) via algorithms remains a chimera. Research efforts within the domain of natural language processing, such as the creation of parallel corpora, text "cleaning" and the morphological and syntactic annotation of texts (pre-editing) have undeniably played a significant role in achieving a level of quality superior to that of rule-based or statistical systems. The results achieved in the field of automatic literary translation, however, do not meet the high demand upheld by human translation (biotranslation). How automatic translation systems are created and how do they function? What parallel corpora are? Are parallel corpora sufficient for generating high-quality automatic literary translations? What quality in translation is? Is there any evaluation metrics for human translations? And what about the metrics used to assess machine translations by humans? What*

* Asist. univ. drd., Facultatea de Litere, Facultatea de Informatică, Universitatea „Alexandru Ioan Cuza” din Iași, Romania; alexandrailie216@gmail.com.

is the current state of automatic translation for texts from classical Romanian literature? These are some of the questions we commit to answering in this paper.

Keywords: *biotranslation; machine translation; corpus; annotation; post-editing.*

Introducere

Lucrarea noastră se situează la confluența dintre literatură și informatică prin însuși obiectul de studiu abordat, traducerea automată. Ne adresăm cercetătorilor din domeniul lingvisticii, al traductologiei, al inteligenței artificiale și al prelucrării limbajului natural, care pot folosi rezultatele studiului pentru a dezvolta teorii sau aplicații; studenților interesați de studierea limbilor străine, a lingvisticii computaționale, a traductologiei, a inteligenței artificiale, care doresc să înțeleagă mai bine evoluția și utilitatea traducerii automate dintr-o perspectivă filologică; traducătorilor interesați de impactul tehnologiilor moderne asupra activității lor; profesioniștilor implicați în crearea sistemelor de traducere și de prelucrare a limbajului natural, dar și tuturor celor interesați și pasionați de traducere și de tehnologia limbajului.

Traducerea s-a situat dintotdeauna la răscruce de drumuri și a oferit condiții prielnice negocierilor și, nu de puține ori, năpastelor. Disprețuită și neîngăduită la începuturi atât în Franța, cât și în România, pe motiv că ar fi atentat la valoarea limbii naționale și pentru că ar fi împiedicat nașterea unei literaturi naționale originale, traducerea nu doar că a supraviețuit, dar a și reușit, în timp, să devină un domeniu de studiu autonom, dezvoltându-și instrumente de lucru și metode proprii, desprinzându-se de alte discipline umaniste centrate pe studierea limbajului. Revoluția tehnologică din ultimii ani nu a exceptat domeniul traducerii. Contactul cu științele exacte (în speță, matematica și informatica) a făcut posibilă apariția unei noi ramuri de cercetare: traducerea automată. Descendentă a traducerii umane, traducerea automată îi împărtășește soarta, fiind, la începuturile sale, la rândul ei, disprețuită și îndelung criticată. Sistemele de traducere prin transfer și cele pe bază de reguli, statistice, rezolvau cu dificultate și cu multe erori sarcina pe care o aveau de îndeplinit. Dar, întocmai modelului primit de la antecesoarea sa, traducerea automată a răzbit cu tenacitate până în zilele noastre, reușind să devină, în contextul globalizării, unul dintre cele mai

utilizate instrumente din era digitală. Cum era firesc, succesele obținute i-au atras și antipatii. Traducătorii umani și-au văzut meseria amenințată, iar în unele cazuri traducerea automată chiar i-a înlocuit pe cei care au considerat cu neputință integrarea sistemelor numerice în desfășurarea activității lor. De altfel, avantajele pe care le oferă sunt incontestabile: rapiditate, costuri reduse, accesibilitate. Dezavantajele nu pot fi excluse din discuție, mai ales atunci când ne referim la traducerea textelor literare: tonul, registrul, efectele stilistice, ambiguitățile textuale, referențialitatea specifice textelor literare sunt dificultăți încă greu de gestionat de sistemele digitale de prelucrare a limbajului natural. Complexitatea acestor tipuri de texte, coroborată cu exigențele impuse de practica traducerii umane, face din traducerea automată un paria. Înainte de a ne decide asupra calității oferite de sistemele automate, propunem o examinare succintă a etapelor și proceselor care contribuie semnificativ la realizarea traducerii automate: colectarea și pregătirea datelor în vederea creării corpusurilor paralele, adnotarea lingvistică, construirea modelelor de limbaj și de traducere, evaluarea calității traducerii.

Pentru că ne-am dorit și o cercetare amănunțită, în cea de-a doua parte a lucrării propunem o analiză comparativă a rezultatelor oferite de sistemele actuale de traducere automată. La baza studiului nostru a stat povestirea *Ion...*¹ de Ion Luca Caragiale. Dat fiind faptul că nu am identificat o traducere umană a acestui text în limba franceză, observațiile noastre se vor mărgini doar la traducerile automate realizate folosind trei dintre cele mai cunoscute sisteme de traducere automată: *Google Translate (GT)*², *Microsoft Translator (MT)*³, *DeepL*⁴, alături de care am inclus și traducerea obținută printr-un sistem nespecializat în această arie, asistentul virtual *Chat GPT*⁵, care include, pe lângă funcționalitățile deja familiare publicului larg, de generare de texte și de interacțiune virtuală, și pe cea de traducere automată. Ne-a interesat să observăm modul în care dificultățile apărute la nivel fonetic, lexical, morfologic, sintactic și stilistic au fost

¹ Ion Luca Caragiale, *Opere III Reminescențe și notițe critice*, ediție îngrijită de Paul Zarifopol, Editura Cultura Națională, București, 1932.

² <https://translate.google.ro/>.

³ <https://translator.microsoft.com/>.

⁴ <https://www.deepl.com/>.

⁵ <https://chatgpt.com>.

depășite de aceste sisteme. Am identificat în textul-suport secvențe și fragmente relevante pentru specificitatea tipului de text și pentru stilul autorului pe care le-am clasificat în funcție de nivelurile limbii. Ținând cont de teoriile apărute în traductologie, am căutat să identificăm care dintre strategiile de traducere transpar în alegerile sistemelor de traducere automată.

În baza rezultatelor obținute, vom formula câteva concluzii referitoare la capacitatea sistemelor avute în vedere de a decoda fenomenele lingvistice investigate și gradul de raportare a lor la teoriile din studiul și practica traducerii umane.

Traducerea automată în 10 pași

Pregătirea textelor pentru procesarea de către un sistem de traducere automat presupune parcurgerea mai multor etape într-o ordine prestabilă. Pentru a observa transformările prin care trece un text literar, de la forma fizică, tipărită pe suport de hârtie (cum este cazul textelor mai vechi), la forma digitală, apoi de la textul sursă la cel țintă, trebuie să descriem, pe scurt, aceste etape.

1. Se impune, întâi de toate, o diferențiere între *digitizare* și *digitalizare*. *Digitizarea* presupune convertirea textului dintr-un format analog într-unul digital și se poate realiza automat prin intermediul sistemelor de *ocerizare*. Dacă, la început, aceste sisteme erau contra cost și îngreunau activitatea cercetătorului, solicitându-l intens pentru corectarea textului, astăzi o simplă fotografiere a textului folosind *Google Lens* este de ajuns pentru recunoașterea grafică a caracterelor și copierea lor într-un computer. În schimb, *digitalizarea* implică transformarea textului digital în formate specifice care pot fi descifrate și gestionate de sisteme numerice de prelucrare a limbajului natural. Cutuma tradițională este stocarea textelor într-un format ușor accesibil, care conține doar informațiile scrise, fără formatare complexă, compatibile cu majoritatea editoarelor de text și ușor convertibile. Pentru aria de studiu ce cuprinde *digital humanities* standardul de formatare este *TEI (Text Encoding Initiative)*, deoarece permite organizarea, stocarea și schimbul de informații între domeniile conexe în mod eficient și interoperabil.

2. Textul brut, obținut prin digitizare trece prin mai multe sub-etape care îi conferă o formă standardizată și simplificată pentru a putea fi procesat de un sistem de calcul. Prin *tokenizare* textul este descompus în unități semantice mici (cuvinte, punctuație etc.) și se obține o listă de *string-uri* (șir de caractere, concept asemănător celui de *sintagmă* în lingvistica generală) grupate într-o singură unitate. Prin *tokenizarea* la nivelul lexical se obține o listă de *tokeni*, iar la nivelul sintactic vom obține o listă de propoziții separate pe baza semnelor de punctuație.

3. Crearea corpusurilor paralele este necesară pentru sarcina specifică a traducerii automate. Mențiunile pe care le vom face în continuare privesc doar perechea de limbi română-franceză și doar tipul de text literar. Dintre colecțiile de texte paralele, libere de drepturi care includ și texte în limba română, menționăm *Opus (Open Parallel Corpus)*⁶, o colecție imensă de corpusuri paralele în continuă dezvoltare. Pentru perechea de limbi română-franceză resursele disponibile cuprind texte din mai multe domenii și sunt în întregime pre-procesate automat, fără intervenție din partea specialiștilor. Acest proiect este dezvoltat și menținut de Universitatea din Uppsala din Suedia și este una dintre cele mai mari și mai utilizate resurse pentru cercetare în domeniul traducerii automate, procesării limbajului natural (NLP) și studiilor lingvistice. Traducerea literară este o sarcină specifică. De aceea, cercetătorii preferă crearea propriilor corpusuri paralele în care să poată include doar acele texte relevante pentru activitatea lor. Dar procesul este unul cronofag și consumator de resurse. Totodată, rezultatele obținute depind în mod direct de calitatea corpusurilor. În teoria traducerii⁷ se insistă asupra identificării și clasificării tipurilor de text al căror scop⁸ traductologic este diferit. În categoria textelor expresive K. Reiß încadrează toate textele elaborate în baza efectelor stilistice: proză literară (eseuri, biografii etc.), proză poetică (anecdote, nuvelă, roman), poezie, baladă etc. Traducerea trebuie să urmărească echivalența la nivelul efectului produs asupra cititorului prin remodelare formală, și nu o imitare servilă (adoptare) a formei din limba sursă.

⁶ <https://opus.nlpl.eu/>.

⁷ Katharina Reiß, *Möglichkeiten und Grenzen der Übersetzungskritik*, Max Hueber Verlag, München, 1986.

⁸ Katharina Reiß, Hans J. Vermeer, *Towards a General Theory of Translational Action: Skopos Theory Explained*, Routledge, London, 2014.

4. Urmează etapa de convertire a majusculilor în minuscule (*lowercase*) și eliminarea punctuației, fapt ce duce, de pildă, la imposibilitatea identificării interjecțiilor. Unul dintre exemplele identificate de noi pe care îl vom prezenta în partea a II-a este cazul interjecției „*Ei!*” care devine omonimă cu forma de pronume personal, persoana a III-a plural, „*ei*” în lipsa punctuației și a majusculei, deoarece traducerea automată se face la nivel de *tokeni*, așa cum am observat mai sus. Așadar, corect este ca rezultatul pe care îl așteptăm pentru traducerea în limba franceză a acestei forme din limba română să fie „*ils*”, echivalentul pronumelui personal de persoana a III-a, plural din franceză, ceea ce se și întâmplă. Din punct de vedere statistic, forma pronominală *ei* apare în mult mai multe contexte extrase din textele introduse în perioada de antrenament al modelelor de limbă. Crearea corpusurilor paralele este o etapă care presupune intervenția specialiștilor. Sistemul „învață” din mult mai multe contexte că echivalentul în limba franceză pentru românescul *ei* este *ils*. Deciziile pe care un sistem le poate lua variază în funcție de tipul de texte care au participat la antrenament. Interjecția „*ei*” apare, cu precădere, în texte literare sau marcate de oralitate. De unde și necesitatea creării de corpusuri paralele. Așadar, sistemele care oferă ca rezultat „*ils*” (fr.) pentru „*ei*” (rom.) funcționează conform regulilor după care au fost create și nu ar putea să își modifice performanțele decât odată cu schimbarea regulilor de antrenament care creează „competența” acestor sisteme. Este cazul sistemelor care încă nu au integrat arhitectura *Transformer*. Mai jos, propunem o analiză comparativă a acestor sisteme. Convertirea majusculilor în minuscule, inclusiv a substantivelor proprii în substantive comune va antrena un nou val de confuzii (substantivul comun *gură-cască* comparativ cu supranumele *Gură-cască*).

5. Sunt eliminate diacriticele, etapă numită *normalizare*. Eliminarea diacriticilor dintr-un text literar atrage un șir de dificultăți care afectează atât acuratețea analizei, cât și calitatea rezultatelor obținute. Cuvintele devin ambigue în lipsa diacriticilor care să le diferențieze (*ăsta* față de *asta*). Analiza morfologică se complică în lipsa posibilității de discriminare a formei de prezent de cea de infinitiv sau de imperfect la persoana a III-a singular (*sta* față de *stă*), sau a nominativului de vocativ (*Ană* față de *Ana*), fapt ce va avea un impact și asupra unei analize stilistice automate, în măsura în care aceasta ar fi posibilă, și, implicit, asupra traducerii finale.

Este important să subliniem faptul că ambiguitățile despre care am amintit se propagă și în etapa de traducere propriu-zisă, de unde și o potențială cauză a rezultatelor obținute de sistemele statistice sau de cele bazate pe rețele neuronale, cum sunt *Google Translate*, *Microsoft Translator* sau, mai nou, *DeepL*. Recent, unele dintre acestea au început integrarea arhitecturilor *Transformer*, reușind să obțină rezultate mai bune. Fiind o descoperire relativ recentă, implementarea suferă modificări, deci și traducerile diferă de la o interogare la alta. Mai mult, modelele noi sunt antrenate pe corpusuri mari de texte paralele, ce pot include toate tipurile de texte.

6. Cuvintele „ne-utile” sunt eliminate. Este vorba despre acele cuvinte frecvente, dar fără un impact major în transmiterea mesajului global și în crearea sensului sau semnificației unui text, printre care amintim: conjuncții, prepoziții, pronume demonstrative, reflexive, articole, unele adverbe sau alte cuvinte irelevante pentru analiza specifică. Acestea sunt înlăturate printr-un *stoplist* specific fiecărei limbi în parte. Scopul este acela de a ajuta sistemul să identifice mai ușor cuvintele relevante cercetării. De identificarea corespunzătoare a „cuvintelor cu cheie” și de analiza lor în conformitate cu întreg conținutul textului depinde, ulterior, întregul proces de traducere, atât în cazul traducerilor umane, cât și al celor automate. Această listă de cuvinte „ne-utile” are rolul de a evidenția cuvintele pertinente în analiza textului. Pentru analiza textelor au fost dezvoltate numeroase instrumente, de la bibliotecile *nlTK* (*Natural Language Toolkit*), *SpaCy*, *Tidytext* (*R package*), până la aplicații *web* sau *desktop*, precum *Voyant Tools*⁹, *Orange Data Mining*¹⁰ ș.a.

7. Unele sisteme pot fi antrenate pe texte care conțin și informație morfologică. Adăugarea acestor date se face prin atribuirea unei etichete fiecărui *token* la nivel de parte de vorbire. Urmează *parsarea*, adică elaborarea unui arbore sintactic în care sunt incluse structurile gramaticale ale propozițiilor unui text. În acest context amintim *UDPipe*¹¹, o resursă bine antrenată pentru limba română (baza de date *UD Romanian Nonstandard* cuprinde texte literare adnotate).

⁹ <https://voyant-tools.org/>.

¹⁰ <https://orangedatamining.com>.

¹¹ <https://lindat.mff.cuni.cz/services/udpipe/>.

8. Alinierea corpusului înseamnă activitatea de așezare a textului original în paralel cu traducerea realizată de traducători umani (biotraducerea, traducerea umană). Unul dintre algoritmi tradiționali de aliniere automată este *Giza++*¹². Acesta se bazează pe modelele IBM utilizate pentru a estima probabilitățile de aliniere a cuvintelor și frazelor între texte. Recent, a fost înlocuit de metode de învățare automată precum modele neurale pentru traducere. Ca și în cazul analizatoarelor de text, există și aplicații web sau desktop de aliniere. Astfel sunt construite corpusurile paralele bilingve sau multilingve, utilizate în traducerea automată și prelucrarea limbajului natural. Procesul de aliniere a textului a fost îmbunătățit de-a lungul timpului. De la alinierea manuală s-a trecut la alinierea automată. Modelele *Transformer* utilizate în arhitecturi precum *BERT*¹³, *XLM-R*¹⁴, *MUSE*¹⁵ aliniaza segmente de text prin învățarea contextuală și calcularea similitudinilor dintre propoziții în spațiul vectorial. Un scor de similaritate ridicat este un indiciu pentru potrivirea între propoziții. *FastAlign*¹⁶ împarte textele în perechi de propoziții și estimează probabilitatea ca fiecare cuvânt din propoziția sursă să corespundă unui cuvânt din propoziția țintă. Tot pe baza similarității lexicale funcționează și

¹² <https://www2.statmt.org/moses/?n=FactoredTraining.RunGIZA>

¹³ Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova, „BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”, în *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Vol. 1 (Long and Short Papers), pp. 4171-4186, Association for Computational Linguistics, Minneapolis, Minnesota, 2019.

¹⁴ Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, Veselin Stoyanov, „Unsupervised Cross-lingual Representation Learning at Scale”, în *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, online, Association for Computational Linguistics, 2020, pp. 8440-8451.

¹⁵ Mimansa Jaiswal, Cristian-Paul Bara, Yuanhang Luo, Mihai Burzo, Rada Mihalcea, and Emily Mower Provost, „MuSE: a Multimodal Dataset of Stressed Emotion”, în *Proceedings of the Twelfth Language Resources and Evaluation Conference*, European Language Resources Association, Marseille, 2020, pp. 1499-1510.

¹⁶ https://github.com/clab/fast_align.

*HunAlign*¹⁷. Acest instrument folosește și un dicționar bilingv de cuvinte. Cu cât alinierea depășește nivelul lexical și se produce pe segmente mai mari de text, cu atât traducerea are șanse mai mari să fie una calitativă. Dar, întrucât nivelul sintactic al limbii este cel mai rezistent la analiză, sarcina de aliniere la nivel sintactic este foarte dificilă. Dificultățile pornesc de la topica frazelor, specifică fiecărei limbi în parte. Un exemplu este distribuția adjectivului față de substantiv în grupul nominal. Dacă în limba română adjectivul este, în general, postpus, în limba engleza sau franceză găsim adjectivele antepuse substantivului. Propoziția subiectivă având ca element regent un verb predicativ se construiește, în general, cu subordonata antepusă regentei. Mai mult, textele literare înregistrează abateri voite de la normă în scopuri stilistice.

9. Toți acești pași se concretizează în antrenarea modelelor de traducere. Memoriile de traducere (*Translation Memories – TMs*) sunt folosite cu precădere în traducerea asistată de calculator (*Computer-Assisted Translation – CAT*) și sunt complementare cu traducerea automată, îmbunătățind rezultatele.

10. După antrenarea modelelor de limbă și de traducere se efectuează traducerea automată. Rezultatele sunt evaluate automat folosind metrici universale precum *BLEU (Bilingual Evaluation Understudy)*¹⁸, *ROUGE (Recall-Oriented Understudy for Gisting Evaluation)*¹⁹, *METEOR (Metric for Evaluation of Translation with Explicit ORdering)*²⁰. Evaluarea automată

¹⁷ D. Varga, P. Haláksy, A. Kornai, V. Nagy, L. Németh, V. Trón, „Parallel corpora for medium density languages”, în *Proceedings of RANLP-05*, Borovets, Bulgaria, 2005, pp. 590-596.

¹⁸ Kishore Papineni, Salim Roukos, Todd Ward, Wei-Jing Zhu, „Bleu: a Method for Automatic Evaluation of Machine Translation”, în *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Philadelphia, Pennsylvania, 2002, pp. 311-318.

¹⁹ Chin-Yew Lin, „ROUGE: A Package for Automatic Evaluation of Summaries”, în *Text Summarization Branches Out*, Association for Computational Linguistics, Barcelona, 2004, pp. 74-81.

²⁰ Michael Denkowski, Alon Lavie, „Meteor Universal: Language Specific Translation Evaluation for Any Target Language”, în *Proceedings of the Ninth Workshop on Statistical Machine Translation*, Association for Computational Linguistics, Baltimore, Maryland, 2014, pp. 376-380.

reprezintă confruntarea textului tradus automat cu o traducere de referință (biotraducere). În lipsa unei traduceri de referință evaluarea revine specialiștilor umani. Evaluarea umană a traducerilor literare nu poate fi standardizată, dar ține cont de strategiile de traducere²¹ și adecvarea lor la tipul de text de tradus. Traducerile pot fi evaluate în funcție de criteriul orientării²², fie traducerea este orientată spre cititor (*naturalizare*²³), fie spre textul sursă (*depeizare*²⁴). O altă abordare²⁵ este cea lingvistică și ține cont de gramatica și de poetica traducerii. Sunt propuse soluții²⁶ pentru fenomene intraductibile. Conceptul de *echivalență*^{27 28 29} între textul sursă și textul țintă

²¹ Jean-Paul Vinay și Jean Darbelnet (*Stylistique comparée du français et de l'anglais. Méthode de la traduction*, Didier 1972) deosebesc strategiile de traducere directă (împrumutul, calcul, traducere literală) de strategiile de traducere oblică (transpoziția, modularea, echivalența, adaptarea, compensația, explicitarea, generalizarea, obscurizarea, concizia, particularizarea, decupajul).

²² Friedrich D. E. Schleiermacher, *Des différentes méthodes du traduire (Ueber die verschiedenen Methoden des Uebersetzens)*, trad. de A. Berman, Éditions du Seuil, Paris, 1999.

²³ *Naturalizarea* (fr. *acclimatation / domestication / naturalisation*; en. *domestication / naturalisation / naturalization*) este o tehnică sau strategie de traducere prin care „traducătorul acordă prioritate fluentei textului restituit în LT și minimalizează, deci, particularitățile identitare ale TS, prin adaptarea culturilor ori a altor mărci culturale la așteptările cititorului-țintă.” (Georgiana Lungu-Badea, *Mic dicționar de termeni utilizați în teoria, practica și didactica traducerii*, ediția a III-a, Editura Universitatii de Vest, Timișoara, 2012, p. 104). În chestiunea privitoare la fidelitate și infidelitate, traducătorul se situează de partea destinatarului, a cititorului.

²⁴ *Depeizarea* (fr. *dépayement / étrangéisation / exotisation*; en. *alienation / exoticization / foreignisation / foreignization*) este o tehnică sau strategie de traducere prin care traducătorul acordă prioritate textului original din limba sursă și constă în transferul unui număr maxim de elemente proprii mediului cultural al textului sursă și are efectul de a face vizibil traducătorul (Stephanie Schwerter, „D'une culture à l'autre. Le défi de traduire les marqueurs régionaux”, în *TTR: traduction, terminologie, rédaction*, Vol. 29, nr. 1, 2016, pp. 115-138). Un echivalent al termenului în limba română ar fi *înstrăinare* (Eugene A. Nida, *Traducerea sensurilor: Traducerea: posibilă și imposibilă*, trad. de Rodica Dimitriu, Institutul European, Iași, 2004, p. 24).

²⁵ George Steiner, *După Babel. Aspecte ale limbii și traducerii*, trad. de Valentin Negoită și Ștefan Avădanei, Editura Univers, București, 1983.

²⁶ Georges Mounin, *Les problèmes théoriques de la traduction*, Gallimard, Paris, 1994.

²⁷ John C. Catford, *A Linguistic Theory of Translation*, Oxford University Press, Oxford, 1965.

orientează evaluarea traducerilor. De asemenea, o altă perspectivă este teoria transparenței. Miza este crearea unei traduceri care să lase impresia că a fost gândită și scrisă în limba țintă³⁰. Conform acestei teorii trebuie investigate strategiile prin care traducătorul (sau chiar sistemul de traducere automată) maschează, elimină, transpune cuvinte, turnuri de frază, tonuri, sentimente, atitudini, moravuri specifice culturii sursă, pe scurt, orice element care ar depeiza și ar dezorienta cititorul. În cazul în care se optează pentru depeizarea textului, sunt investigate procesele prin care elementele și fenomenele specifice textului și culturii sursă ajung să fie transpuse în textul țintă. În acest caz, evaluarea se face prin observarea alegerilor făcute de traducător. Acestea pot include transferul *verbatim* al cuvintelor străine, păstrarea structurii gramaticale a limbii dinspre care se traduce sau chiar inserarea arhaismelor cu efect de reconstituire a izului unui secol apus, în cazul textelor vechi³¹. Mounin atrage atenția asupra efectului de disparitate creat în traduceri, cauza principală a eterogenității textului țintă fiind lipsa de uniformitate lingvistică, stilistică, de ton, de registru etc.³² L. Venuti critică tendința himerică de creare a traducerilor „transparente”³³, deoarece ar avea ca efect subaprecierea eforturilor și a intensității travaliului prin care traducătorul este nevoit să treacă, mai ales în cazul traducerilor literare. În teoria skoposului³⁴, criteriul de evaluare rămâne echivalența. *Contextualizarea* este un concept central pe care se sprijină noile cercetări în domeniul traductologiei. Toury³⁵ evidențiază

²⁸ Eugene Nida, Charles R. Taber, *The Theory and Practice of Translation*, Vol. III, ed. II, E. J. Brill, Leiden, 1982.

²⁹ Mona Baker, *In other words: a coursebook on translation*, Routledge, London – New York, 1992.

³⁰ Georges Mounin, *Les belles infidèles*, Presses Universitaires de Lille, Paris, 1994, p. 74.

³¹ Georges Mounin, *op. cit.*, p. 92.

³² Jean Delisle, „L'évaluation des traductions par l'historien”, *Meta Journal des traducteurs*, Vol. 46, nr. 2, 2001, pp. 209-226.

³³ Lawrence Venuti, *The Translator's Invisibility. A History of Translation*, ed. a II-a, Routledge, Oxon – New York, 2008, p. 1: *The illusion of transparency is an effect of a fluent translation strategy, of the translator's effort to insure easy readability by adhering to current usage, maintaining continuous syntax, fixing a precise meaning.*

³⁴ Katharina Reiß, Hans J. Vermeer, *op. cit.*, 2014.

³⁵ Gideon Toury, *Descriptive Translation Studies – and beyond*, John Benjamins Publishing Company, Amsterdam-Philadelphia, 2012.

perspectivele multiple de analiză a traducerii, în vreme ce Itamar Even-Zohar³⁶ propune conceptul de *polisistem* pentru a descrie fenomenul de interconectare a unor ansambluri de sisteme responsabile de apariția textului sursă și de generarea unui text țintă corespondent.

În continuare propunem o evaluare critică și comparativă a traducerilor în limba franceză a povestirii *Ion...* de Ion Luca Caragiale. Vom analiza fenomenele de natură fonetică, lexicală, morfologică, sintactică și stilistică. Vom stabili ce strategii de traducere pot fi decelate din rezultatele obținute. Pentru că nu am găsit publicată o traducere umană, care să devină traducerea de referință pe baza căreia se calculează scor *BLEU*, nu am putut include în acest studiu și evaluarea automată a traducerilor automate realizate.

Evaluarea traducerilor automate. Studiu de caz: *Ion...* (Caragiale) în limba franceză

În continuare extragem câteva secvențe de text care ar putea genera dificultăți de traducere. Observăm comparativ modul în care *Google Translate*, *Microsoft Translator*, *DeepL* și *Chat GPT* au depășit obstacolele identificate la nivel fonetic, lexical, morfologic, sintactic și stilistic. Stabilim în ce măsură putem confirma prezența strategiilor de traducere tradiționale.

Fenomene de natură fonetică:

Ion Luca Caragiale	Google Translate	Microsoft Translator	DeepL	Chat GPT
„pân' la sânge”	„jusqu'au sang”	„jusqu'au sang”	„sans cesse”	„constamment”
„- Ei! și pe urmă?”	„- Ils! et alors?”	„- Salut! Et puis?”	„- Hé! Et après?”	„- Eh bien! et après?”
„ Ei! vă place?...”	„Ey! aimez-vous?”	„Son! Vous l'aimez?...”	„Eh bien! ça vous plaît?”	„Alors? Ça vous plaît?...”
„- De, mă băietele!”	„- Oui, moi garçon!”	„- Oui, moi garçon!”	„- De, mon garçon!”	„- Eh bien, mon garçon!”
„- Ei! uite așa”	„- Ils! Regarder comme ça”	„- Salut! Voici comment”	„- Et voilà!”	„- Eh bien! comme ça”

³⁶ Itamar Even-Zohar, „Poetics Today. Polysystem Studies”, în *International Journal for Theory and Analysis of Literature and Communication*, Vol. XI, Duke University Press, Tel Aviv, 1990.

Traducerea automată a literaturii. O himeră încă vie?

„măgar — din gură Carnavalul de Venezia—cu variațiuni?... Ai?... ”	„âne — de la bouche Carnaval de Venise — avec des variantes?... Et vous?... ”	„âne, de la bouche du carnaval de Venise, avec des variantes?... Ail?... ”	„l'âne — de sa bouche Le Carnaval de Venise — avec des variations?... Vous avez?... ”	„Un âne — avec sa bouche, Le Carnaval de Venise — avec des variations!? Hein?... ”
„- Ia, ascultă, mă băiețe”	„- Allez , écoute - moi garçon”	„- Écoutez, mon garçon”	„- Écoute, mon garçon,”	„-Écoute- moi , jeune homme”
„- Nu se poate, mă omule”	„- Ce n'est pas possible , maman homme”	„- Il n'est pas possible, homme”	„- Il n'est pas possible, mon homme”	„- Ce n'est pas possible, mon bon homme”
„Că nu vrei tu, ai? Ei! iaca , la toți ne place, să crăpi! ”	„Que tu ne veux pas as- tu Ils! oui , nous l'aimons tous , oui fissure ! ”	„«Tu ne veux pas, n'est-ce pas?» Son! Eh bien, nous <u>aimons tous</u> crâquer ! ”	„Vous ne voulez pas, n'est-ce pas? Eh bien, <u>nous aimons</u> <u>tous mourir !</u> ”	„Parce que tu ne veux pas, hein? Eh bien, tu vois , ça nous plaît à tous, crève ! ”
„- Niște ce, mă? ”	„- Quelques quoi, moi? ”	„- Quoi, mec? ”	„- Quelques-uns quoi, mon pote? ”	„- Des quoi, hein? ”
„- Ei! du-te la dracul cu prostiile tale!”	„ - Ils! au diable vos bêtises !”	„ - Salut ! Allez au diable avec vos bêtises!”	„ Eh bien , allez au diable avec vos couilles!”	„ - Eh bien ! va au diable avec tes bêtises!”

În ceea ce privește fenomenele de natură fonetică, am limitat selecția la interjecții și semne de punctuație. Prezența interjecțiilor „*Ei!*”, „*Uite!*”, „*Iaca!*”, „*mă!*”, „*Ai!*”, „*De!*” creează o atmosferă dinamică, conferind un ritm alert textului. Acestea contribuie și la crearea unui limbaj colocvial și natural. De asemenea, folosirea interjecțiilor scoate în evidență momentele culminante sau critice din poveste. Inserarea interjecțiilor în dialogul personajelor apropie textul literar avut în vedere de latura vorbită a limbii și conferă stilului o valoare de oralitate. Fiind un maestru al vorbelor, Caragiale cunoaște foarte bine valoarea cuvintelor pe care le întrebunțează. Nicio alegere nu este întâmplătoare. În *Ion...* interjecția „*Ei!*” apare de cinci ori, de două ori cu rolul de a întări o întrebare („și pe urmă?” și „vă place?”) prin care se face și trecerea de la o idee la alta, de două ori însoțită de o altă interjecție, cu rolul de a întări o mustrare („uite așa”, „iaca”) și o dată cu rolul de a exprima o poruncă („du-te

la dracul cu prostiile tale!"). Soluția GT pentru patru din cele cinci cazuri este traducerea interjecției prin forma de pronume personal, persoana a III-a, plural. În seturile de date de antrenament ale GT, este posibil ca „Ei!” să fi fost asociat mai des cu situații în care apare aliniat cu forma pronominală, rezultând în această traducere. O singură dată GT optează pentru traducerea prin „Ey!”. Motivul pentru care aici nu am regăsit forma „Ils” este contextul în care interjecția apare în textul original „bravo! și bis!... ..Ei! vă place?...”, mai exact, însoțit de alte forme exclamative. Însă forma aceasta „ey” nu este o interjecție normată a limbii franceze, cel mult ar putea fi un morfem care apare în unele cuvinte franțuzești pentru a explica un è protonic, ca în *asseyons*. Pentru cele cinci ocurențe în limba română MT traduce interjecția prin „Salut!” (X 3) și „Son!” (X 2). Apariția formulei de salut și a adjectivului posesiv de persoana a III-a, fr. „son” ca echivalent pentru forma de genitiv feminin singular, ro. „ei”, se poate explica prin faptul că MT folosește modele statistice care se bazează pe un set mare de date de antrenament. Calculând frecvența, forma „salut” a obținut cel mai mare scor și a devenit echivalentul formei grafice „Ei!”, convertită prin *tokenizare și lowercase* la „ei” în noua sarcină de traducere. DeepL a tradus cele cinci ocurențe din limba română prin „Hé!”, „Eh bien!”, „Et voilà!”. „Hé!” este o interjecție cu rol de exprimare a mirării sau pentru a formula un salut. Contextul în care apare este unul care exprimă chiar mirarea. „Eh bien” are rolul de a întări o interogație, „Eh bien ! ça vous plaît?”, sau de a exprima indignarea „Eh bien, nous aimons tous mourir!” (!). Chat GPT traduce de patru ori prin „Eh bien!” și o dată prin adverbul „Alors!”, în contextul în care apare și pronumele *Ça* „...Alors ? Ça vous plaît?...”, deoarece registrul familial al limbii franceze înregistrează această posibilitate a asocierii cu un pronume pentru a atribui discursului o nuanță de surpriză sau de indignare. Însă în etapa de recuperare a punctuației adverbul *alors* este privat de această nuanță, dar prezența sa în acest context capătă sens concludiv. În textul sursă interjecția „Ei!” din contextul „Ei! vă place?...” apare cu scopul de a solicita o opinie din partea ascultătorilor și cere încheierea șirului povestirii cu o concluzie. Interjecția „mă!” apare de patru ori în dialogul lui Ion cu oamenii de la han, de trei ori enunțată de alocutorii lui Ion, o dată de Ion. Cuvântul acesta de adresare, fără funcție sintactică, marchează între vorbitori un raport de la egal la egal sau de la superior la inferior. Traducerea s-a făcut fie printr-un adjectiv posesiv (*mon*; X 7) cu funcția de atribut, fie printr-un pronume în acuzativ (*moi*; X 5) cu funcția de

complement direct. Motivația alegerii este omonimia dintre interjecția *mă* și pronumele personal de persoana I singular în acuzativ. Dezambiguizarea se poate face doar din context. *GT* găsește ca soluție și forma de vocativ „*mamam*” (*mamă*), *MT* a tradus o dată printr-un termen argotic „*mec*” (băiat), *DeepL* o dată printr-un termen din registrul familiar „*pote*” (amic). *Chat GPT* aduce și soluția de traducere prin interjecția „*hein!*” cu rol în exprimarea indignării în contextul „- *Des quoi, hein?*”, singura soluție pe care o considerăm potrivită, mai aproape de intențiile autorului. „*Hein*” recuperează valoarea interogativ-exclamativă a enunțului și exprimarea indignării, dar periclitează relația pe care o menține „*mă*” în adresare. Există o diferență de nuanță între interjecția onomatopeică „*hein*” și interjecția articulată „*mă*”. O altă soluție ar fi interjecția fr. „*wesh*” utilizată în limbajul colocvial sau familial cu rol de interpelare. Îndeplinește, la fel ca *mă*, funcție fatică de stabilire și de păstrare a comunicării și a relației dintre emițător și receptor. Alte soluții ar fi utilizarea unui substantiv în vocativ precum „*pote*” (așa cum am observat în cazul *MT*) sau chiar „*frérot*”. Observăm în toate cazurile o alternanță în decizii. Lipsa de uniformitate în traducerea unei forme unice, impusă de autor, prin două sau chiar trei variante are un impact negativ asupra coerenței textului țintă. Soluțiile oferite evidențiază provocările cu care se confruntă aceste sisteme de traducere automată atunci când trebuie să gestioneze interjecții. Interpretarea nuanțelor subtile este o sarcină încă dificilă pentru sistemele care analizează textul literal, fără a lua în considerare contextul cultural sau emoțional.

Fenomene de natură lexicală

Ion Luca Caragiale	Google Translate	Microsoft Translator	DeepL	Chat GPT
„un zăduf de nămiez”	„une cabane en rondins”	„une masse”	„Sur un champ de neige”	„une chaleur étouffante de midi”
„un flăcău, drumeț sărac, ostenit și prăfuit”	„un garçon arrêté, un randonneur pauvre, fatigué et poussiéreux.”	„un jeune homme s'est arrêté, un pauvre randonneur fatigué et poussiéreux.”	„un garçon, un pauvre voyageur fatigué et poussiéreux”	„un jeune homme s'arrêta, un pauvre voyageur, fatigué et couvert de poussière”

„s-a așezat cuviincios pe o laviță ,”	„je me suis assis correctement sur un banc ”	„s'assit consciencieusement sur un lavabo ”	„s'assit de bonne manière sur une petite table ”	„s'assit poliment sur un banc ”
„o bărdacă de vin mai ieftior .”	„une carafe de vin pas cher ”	„un tonneau de vin ”	„un tonneau du meilleur vin ”	„une chope de vin bon marché ”
„cu ce rost și daraveri ”	„avec quoi articulation et daraveri ”	„à quoi cela sert ”	„quel est son but et ses antécédents ”	„ quelles étaient ses affaires ”
„că soră-sa a avut noroc ; iar el, ba: ea a ajuns procopsită, doamnă mare, primită pe la curți împărătești ”	„que sa sœur avait chance et lui, ba: elle est arrivée procopsita, grande dame, reçue au tribunal tu es roi ”	„que sa sœur avait de la chance; Et lui, bien: elle est devenue une procopsita, une grande dame, reçue à la cour royale ”	„que sa sœur avait de la chance, et que lui n'en avait pas: Elle est devenue une grande dame, une grande dame, une grande dame, une dame des cours royales; ”	„Que sa sœur avait eu de la chance, tandis que lui, non: elle était devenue prospère, une grande dame accueillie dans les cours royales ”
„nu te poți tu, băiat voinic, sălta din nevoie ”	„pas toi tu peux toi mon garçon volontaire, saute par nécessité ”	„tu ne peux pas, garçon fort, sauter par nécessité ”	„que tu ne peux pas, toi, le costaud, te libérer de ton besoin. ”	„ne pas réussir, toi, un jeune homme robuste, à sortir de la misère. ”
„halal să le fie!... așa mângâietură, n-am ce zice ”	„que ce soit halal pour eux!... alors réconfort, je n'ai rien il dit ”	„pas question... halal... une telle caresse, je n'ai rien à dire ”	„«Malheur à eux!»... une telle caresse, je n'en dirai jamais assez »	„bravo à eux!... je n'ai rien à dire, une telle correction ”
„La o răspântie, unde se strânsese multă gură-cască, un pristav, cu o chivără mare de hârtie ”	„Sur une croix, où il s'était rassemblé PLUS un casque buccal, un pristav, avec une grande feuille de papier ”	„À un carrefour, où beaucoup de casques buccaux s'étaient rassemblés, un pristav, avec une grande cruche en papier ”	„A un carrefour, où s'étaient rassemblés de nombreux commérages, un pristav, muni d'un grand chivaree en papier ”	„À un carrefour, où s'était rassemblée une grande foule de badauds, un héraut, coiffé d'un énorme chapeau de papier, ”

Traducerea automată a literaturii. O himeră încă vie?

„Grădina era luminată cu multe lumini în bășiți de hârtie vâpsită ”	„Le jardin était bien éclairé lumières dans bulles de papier tein ”	„Le jardin était éclairé de nombreuses lumières dans des cloques de papier verni. ”	„Le jardin était éclairé par de nombreuses lumières dans des mèches de papier voilé. ”	„Le jardin était illuminé par de nombreuses lumières dans des lanternes de papier coloré. ”
„Ion a-nceput să râză de mofturile caraghiozului ”	„Ion commença selle rire des modes caraghioz ”	„John se mit à rire des caprices du farceur ”	„Ion se mit à rire des pitreries de cet imbécile ”	„Ion s’est mis à rire des caprices du charlatan ”
„am găsit multă lume adunată, și din boierime și din tagme și prostime – destulă ”	„j’ai trouvé PLUS monde collectés et de la noblesse et de tagme moi et absurdité – ça suffit. ”	„je trouvai beaucoup de gens rassemblés, tant des boyards que des Tagme et des Nonsenses – assez. ”	„j’ai trouvé beaucoup de gens rassemblés, à la fois des nobles et des gens du peuple - pas mal de monde. ”	„j’ai trouvé une foule considérable, composée de nobles , de bourgeois et de gens du peuple – il y en avait beaucoup ”
„îi sclikeau capul, pieptul și mâinile de diamantice. ”	„EUX la tête, la poitrine brillaient et mains en diamant ”	„Sa tête, sa poitrine et ses mains brillaient comme des diamants. ”	„sa tête, sa poitrine et ses mains scintillaient de diamants. ”	„sa tête, sa poitrine et ses mains brillaient de diamants. ”
„Se arătase sus pe movilă vestitul cântăreț, gătit cu un valdrap de mătase cusut numa-n fireturi , și decorații atârnați de gât, și fel de fel de zorzoane, ”	„Il s’était montré sur le monticule célèbre chanteur, cuisiné avec un valdrap de soie couture numa-n threads , et PRIX pendu au cou, et sorte de zorzoane ”	„Le célèbre chanteur était apparu sur le monticule , cuit avec un drap de soie cousu seulement en fils , et des décorations pendaient à son cou, et toutes sortes de fioritures ”	„Sur la butte , le célèbre chanteur était apparu, vêtu d’un valdrap de soie cousu de fils seulement , avec des décorations accrochées à son cou, et toutes sortes de fanfreluches, ”	„Le célèbre chanteur était apparu en haut de la colline , vêtu d’un manteau de soie orné de galons dorés , des décorations pendues à son cou, et des colifichets de toutes sortes”

În ceea ce privește alegerile lexicale pe care Caragiale le-a făcut în construirea textului său, observăm predilecția pentru turcisme, grecisme, slavisme și pentru folosirea termenilor cu sens dublu, cel figurat dublând mesajul de suprafață. Ca și cum ar exista două lumi discursive întrepătrunse, una de suprafață vizibilă tuturor, și una de adâncime care se deschide doar în urma unei analize profunde, traducerea oferă prilejul decodării mesajelor încifrate în forma fixă a cuvintelor. La acest nivel, capcana în care pot ușor să cadă traducătorii este utilizarea în exces a procedurii de traducere numit *explicitare*³⁷. Din dorința de a recupera cât mai mult din semnificația unui termen, abordarea acestei strategii răpește din plăcerea descoperirii textului și a semnificațiilor ascunse de către cititorul însuși.

Să luăm, întâi de toate, exemplul „*zăduf de nămiez*” care a fost tradus prin „*cabane en rodins*” (cabană din bușteni) (GT), „*masse*” (masă) (MT), „*champ de neige*” (DeepL) (câmp de zăpadă) și „*chaleur étouffante de midi*” (căldură sufocantă de amiază) (ChatGPT). Cuvântul *zăduf* are în limba noastră o semnificație dublă. Cu sens propriu înseamnă „căldură mare, înăbușitoare”, iar cu sens figurat „supărare, necaz”, potrivit DEX, 2016. Este un derivat al verbului *zăduși*, provenind din sl. *zadušiti*³⁸. Dintre toate sistemele, doar *Chat GPT* a păstrat una dintre semnificații, aceea de *căldură sufocantă, înăbușitoare de amiază*. Trimiterea sugestivă la *necaz*, precum cel despre care avea să povestească Ion că-l are cu sora lui, pe care o antrenează sensul denotativ al lui *zăduf* s-a pierdut în traducere. Problemele puse de termeni arhaici, regionalisme sau termeni creați de autor, precum „*daraveri*”, „*procopsită*”, „*pristav*”, „*diamanticală*”, „*valdrap*”, au fost rezolvate prin transfer direct (împrumut) de GT și MT. „*Daraveri*” are, de asemenea, pe lângă sensul de *treabă, afacere comercială*, și pe cel de *încurcătură*. Etimologia termenului este încă nestabilă: fie provine din tc. *dalavera* (înșelătorie), fie

³⁷ Georgiana Lungu-Badea, *Mic dicționar de termeni utilizați în teoria, practica și didactica traducerii*, ed. a III-a, Editura Universitatii de Vest, Timișoara, 2012, p. 72: „utilizată atunci când împrumutul, calcul sau echivalența nu ajută la actualizarea semnificației intenționate de autor”.

³⁸ Alexandru Ciorănescu, *Dicționarul etimologic al limbii române*. Ediție îngrijită și traducere din limba spaniolă de Tudora Șandru Mehedinți și Magdalena Popescu Marin, București, Editura SAECULUM I. O., 2001, p. 848.

este un cuvânt compus între *dare* și *avere*³⁹. *DeepL* a optat pentru traducerea prin „*antécédents*” (în sensul de istorie personală), în timp ce *Chat GPT* prin „*affaires*” (treburi). Niciuna dintre opțiuni nu recuperează ambele sensuri pe care *daraveri* le încapsulează. Adjectivul „*procopsită*” provine din participiul verbului *a procopsi*, variantă pentru *a pricopsi*, cu trei sensuri în limba română: a educa (1), a îmbogăți (2) și, cu forma reflexivă, a profita (3)⁴⁰. Provine din ngr. *προκόπτω*. *DeepL* am putea spune că reține doar primul sens prin „*grande dame*”, iar *ChatGPT* doar pe al doilea „*prospère*”. Un fenomen aparte s-a produs la traducerea locului pe care Ion l-a ales: el „*s-a așezat cuviincios pe o laviță*”, ca să-și mai poată trage sufletul. Locurile și maniera în care ar trebui să se așeze, dacă textul ar fi fost tradus prin sisteme automate, ar fi dat naștere unei scene hazlii. Iată și de ce: *GT* traduce tabloul prin „*assis correctement sur un banc*” (implicând faptul că există și o modalitate incorectă), apoi *MT* ne indică faptul că Ion „*s’assit consciencieusement sur un lavabo*”. Oare cât de scrupulos ar trebui să fie personajul ca să se poată așeza pe o chiuvetă? *DeepL* ni-l înfățișează pe Ion în ipostaza în care se așază politicoș pe o măsuță mică „*s’assit de bonne manière sur une petite table*”. Un alt moment, cel al comenzii pe care o face Ion imediat ce ajunge la han, a dat naștere în traduceri automate unor realități distorsionate din textul sursă în care cere „o bărdacă de vin mai ieftior”. *MT* propune o soluție mai generoasă și oferă „*un tonneau de vin*” (un butoi de vin), iar în versiunea *DeepL* se modifică și calitatea vinului „*un tonneau du meilleur vin*” (un butoi din cel mai bun vin). Varianta *GT* „*carafe*” nu recuperează izul arhaic impus de „*bărdacă*” (din tc. *bardák*), dar *Chat GPT* reușește acest lucru prin „*chope*”. *Chat GPT* reușește să traducă acele cuvinte pe care celelalte sisteme le-au împrumutat pur și simplu, dar, chiar și prin efortul de explicitare, nu reușește să recupereze sensurile conotative ale termenilor.

³⁹ *Ibidem*, p. 279.

⁴⁰ *Ibidem*, p. 639.

Fenomene de natură morfologică

Ion Luca Caragiale	Google Translate	Microsoft Translator	DeepL	Chat GPT
„poate că și tu <i>ăi fi vinovat!</i> <i>ăi fi având</i> și tu cine știe ce cusur or slăbiciune”	„peut être que et toi <i>tu es coupable!</i> tu aurais et toi qui sais Quoi faute ou faiblesse”	„Peut-être <i>êtes-vous</i> coupable aussi! Tu peux avoir je ne sais quel défaut ou quelle faiblesse,”	„Peut-être <i>as-tu</i> un défaut ou une faiblesse”	„peut-être que <i>tu es</i> aussi fautif! Tu <i>dois</i> bien <i>avoir</i> un défaut ou une faiblesse”
„Poate <i>ăi fi vinovat!</i> <i>ăi fi având</i> și vrun cusur, ce să zic?”	„Peut-être <i>qu’il aura</i> et courir faute, quoi selle Je dis ?”	„Peut-être que j’ai quelques défauts, que puis-je dire?”	„ <i>J’ai</i> peut-être des défauts, que voulez-vous?”	„Peut-être <i>ai-je</i> un défaut, que dire?.”
„Unde <i>te-i duce</i> , tot de mine <i>ai să dai,</i>	„Où <i>va-t’en</i> , tu m’ <i>auras</i> toujours <i>allez</i> ”	„Là où il vous <i>emmène</i> , vous me <i>trouvez</i> toujours.”	„Là où ils <i>t’emmèneront</i> , tu me trouveras.”	„Où que tu <i>ailles</i> , tu <i>tomberas</i> toujours sur moi.”
„ <i>Oi fi umblând</i> , tu iute, că zbori pe sus cu falaitar, și eu încet, că mă târâsc pe jos; dar tot trebuie să ne-ntâlnim!”	„ <i>Tu marcheras</i> vite, ça envole - toi avec le Falaitar, et je lentement, ça moi ramper sur le sol; mais tu dois quand même rencontrons - nous!”	„ <i>Je marcherai</i> , tu seras rapide, parce que tu voles haut avec un falaitar, et je serai lent, parce que je rampe sur le sol;”	„ <i>Vous avez beau vous promener</i> , vous vite, en volant dans les airs, et moi lentement, en rampant à pied, il faut bien que nous nous rencontrions!”	„ <i>Toi, tu cours vite</i> , tu voles dans les airs avec ta calèche, et moi, je traîne à terre;”

Am discutat în această secțiune doar despre formele de viitor popular și despre pronumele personal. Viitorul popular este mod caracteristic al oralității stilului prin care Caragiale inserează forme caracteristice limbajului oral sau colocvial, cu scopul de a reda cât mai fidel particularitățile registrului familial. Sistemele de traducere nu surprind această caracteristică orală. În franceză, viitorul apropiat este mult mai des utilizat în vorbire decât viitorul standard. În nicio situație de traducere a viitorului popular nu am identificat o formă de viitor

apropiat în franceză. În schimb, acesta a fost redat prin prezent, viitor standard sau subjonctiv. Pronumele personal, persoana a II-a singular a fost tradus de MT și prin pronume personal, persoana a II-a plural, forma de politețe „*Peut-être êtes-vous coupable aussi*”, „*Vous avez beau vous promener, vous vite*” sau prin pronume personal, persoana I singular „*Je marcherai*”.

Fenomene de natură sintactică

Ion Luca Caragiale	Google Translate	Microsoft Translator	DeepL	Chat GPT
„nu fac pas, să nu dau de ea, tot de ea și iar de ea”	„Et je ne fais pas un pas, pour ne pas la toucher, aussi pour elle et par elle.”	„Et je ne fais pas un pas, pour ne pas la croiser, elle aussi et encore elle.”	„Et je ne ferai pas un pas, de peur de la rencontrer, et elle, et elle, et elle.”	„Et je ne fais pas un pas sans tomber sur elle, toujours elle et encore elle.”
„știu asemenea că orce școală, cât de înaltă, nu poate da ce n-a vrut să dea nașterea”	„et je sais aussi que N'IMPORTE LEQUEL L'école, aussi élevée soit-elle, ne peut pas donner ce qu'elle ne veut pas selle donner la naissance.”	„je sais aussi que toute école, aussi haute soit-elle, ne peut pas donner ce qu'elle n'a pas voulu enfanter..”	„je sais aussi qu'aucune école, aussi élevée soit-elle, ne peut donner ce qu'elle n'a pas voulu donner...”	„je sais aussi qu'aucune école, aussi prestigieuse soit-elle, ne peut donner ce que la naissance a refusé...”
„Tuturilor nouă, și împăratului și-mpărătesei, la toți boierii, la tot norodul, mie și dv., ne place la toți”	„À tout le monde neuf, et Empereur et l'impératrice, à tous les boyards, à tout le monde, à moi et toi, on t'aime tous”	„Nous tous, l'empereur et l'impératrice, tous les boyards, tout le peuple, vous et moi, comme nous tous”	„Nous tous, l'Empereur, l'Impératrice, l'Impératrice, tous les nobles, tout le peuple, vous et moi, nous l'aimons tous”	„À nous tous, à l'empereur et à l'impératrice, à tous les nobles, à tout le peuple, à moi et à vous, ça plaît à tout le monde”
„scoțându-și de la mână o brățară, care scânteia de departe, i-a petrecut-o pe dup-o ureche.”	„prenant de sa main un bracelet qui brillait de loin, elle le mit derrière son oreille”	„prenant de sa main un bracelet qui brillait de loin, le passa sur son oreille”	„sortant de sa main un bracelet qui scintillait de loin, elle le mit derrière une de ses oreille”	„en retirant de son poignet un bracelet qui scintillait au loin, elle le lui a passé autour d'une oreille.”

În cazul primului exemplu observăm că fraza în limba română are o structură simplă, creată repetitiv cu scopul de a sublinia neplăcerea fratelui la întâlnirea cu sora sa. *GT* include, în mod eronat, de două ori consecutiv conjuncția coordonatoare copulativă *și*, iar ritmul frazei este întrerupt. „*Par elle*” (prin ea) se îndepărtează de sensul din textul țintă. Repetiția „*elle aussi et encore elle*” recuperează intenția din textul sursă. Forma de viitor a verbului în varianta *DeepL* modifică timpul acțiunii, iar inserarea expresiei „*de peur de*” este inutilă și nepotrivită în context, deoarece sugerează teama personajului; or, în acest caz e vorba despre frustrarea sa și inevitabilitatea întâlnirii. Există ideea de repetiție, dar modul de creare al acesteia are un efect negativ asupra ritmului frazei. *ChatGPT* oferă o traducere fidelă în raport cu textul sursă. Un aspect sintactic important este exprimarea în limba franceză a unui subiect care în limba română este inclus. Motivul ține strict de natura sintactică a limbilor: franceza nu permite flexiunea verbelor fără pronume decât la imperativ. Așadar, sistemele de traducere automată au rezolvat această sarcină corect incluzând un cuvânt care, altfel, nu ar fi avut nevoie de complinire. Complementele indirecte din exemplele „*Tuturilor nouă, și împăratului și-mpărătesei, la toți boierii, la tot norodul, mie și dv., ne place la toți*” au devenit subiecte în versiunile *ML* „*Nous tous, l'empereur et l'impératrice, tous les boyards, tout le peuple, vous et moi, comme nous tous*” și *DeepL* „*Nous tous, l'Empereur, l'Impératrice, l'Impératrice, tous les nobles, tout le peuple, vous et moi, nous l'aimons tous*”.

Tehnici de traducere identificate

La nivel fonetic am observat că *GT* și *MT* se bazează în continuare pe modele statistice, calculul frecvenței ocurențelor, de unde și o traducere, în linii generale, asemănătoare cu traducerile în care s-a abordat una dintre strategiile directe, traducerea literară. Am identificat același lucru, dar într-o măsură mai mică, și la *DeepL*. *ChatGPT* a fost singurul sistem care a reușit să găsească soluții mulțumitoare care să răspundă exigențelor de creare a unei echivalențe între texte și de orientare a textului către cititor.

La nivel lexical *GT*, *MT* și *DeepL* soluționează problema arhaismelor, a regionalismelor, a cuvintelor plurisemantice sau a cuvintelor inventate de autor printr-o altă strategie de traducere directă, împrumutul, procedeu ce constă în transferul direct din limba sursă în limba țintă fără traducerea

cuvântului. Pe de altă parte, *Chat GPT* creează traduceri prin amplificare (subramură a transpoziției, prin care în limba țintă sunt folosite mai multe cuvinte decât în limba sursă pentru exprimarea aceluiși mesaj), explicitare și generalizare (de exemplu, „*prostime*” tradus prin „*gens du peuple*”). Aceste strategii de traducere depășesc traducerile directe, fiind încadrate în categoria traducerilor oblice. Și la acest nivel orientarea traducerii se face dinspre textul sursă spre cititor.

La nivel morfologic observăm că valorile modurilor și timpurilor verbale nu sunt respectate. În cazul traducerii pronumelui se înregistrează confuzii și ambiguități. În general, aceste fenomene sunt traduse direct.

La nivel sintactic am identificat apariția *expansiunii* prin faptul că structurile sintactice echivalente includ pe lângă elementele regente determinanți prepoziționali. Limbile au o organizare sintactică proprie. Subiectele reluate și cele incluse din textul sursă au avut nevoie de transformări în textul țintă. În linii generale, sistemele de traducere automată îndeplinesc această cerință prin faptul că antrenarea nu s-a făcut doar la nivel de cuvânt, ci și de sintagmă.

Concluzii

În studiul de față au fost aduse în discuție diferite perspective de abordare a traducerii. Am coroborat studiile de traductologie cu cele de lingvistică computațională și de prelucrarea limbajului natural. Lucrarea cuprinde și o secțiune în care sunt prezentate 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 corpusuri paralele, adnotarea, antrenarea modelelor de limbă și de traducere, evaluarea. Am evaluat traducerile realizate de sistemele automate *Google Translate*, *Microsoft Translator*, *DeepL* și *Chat GPT* după criteriile impuse de cercetările traductologice făcând un studiu de caz pe textul literar *Ion...* de Ion Luca Caragiale. Am căutat să observăm soluțiile oferite de aceste sisteme pentru dificultăți apărute la nivel fonetic, lexical, morfologic, sintactic și stilistic și am identificat strategiile de traducere apărute în variantele traduse.

Traducerea automată a textelor literare este încă într-o etapă de tatonare. Modele de generație anterioară funcționează în paralel cu cele de generație mai nouă. Cazul *DeepL* este sugestiv în acest sens.

Abordarea unei tehnici de traducere directă poate duce la crearea unui „dezechilibru” cantitativ între textul sursă și textul țintă. *GT* și *MT* se bazează în continuare pe strategii directe. Totodată, am identificat în variantele *Chat GPT* și *DeepL* strategii de traducere oblică precum explicitarea, transpoziția, expansiunea, generalizarea.

Noile arhitecturi de inteligență computațională au o rată mai mare de succes, însă depind în continuare de post-editare din partea specialiștilor. Tendința de adăugare a informațiilor care în textul sursă doar au fost sugerate, lăsate intenționat în acea formă de către autor spre a fi identificate și interpretate de cititor, răpește din plăcerea lecturii. Cititorul primește mesajul transmis de autor gata interpretat, uneori incorect sau insuficient.

Bibliografie

- Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, Veselin Stoyanov, „Unsupervised Cross-lingual Representation Learning at Scale”, în *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, online, Association for Computational Linguistics, 2020, pp. 8440-8451.
- Baker, M., *In other words: A coursebook on translation*, Routledge, London – New York, 1992.
- Caragiale, I. L., *Opere III, Reminiscențe și notițe critice*, ediție îngrijită de Paul Zarifopol, Editura Cultura Națională, București, 1932.
- Catford, J. C., *A Linguistic Theory of Translation*, Oxford University Press, Oxford, 1965.
- Alexandru Ciorănescu, *Dicționarul etimologic al limbii române*. Ediție îngrijită și traducere din limba spaniolă de Tudora Șandru Mehedinți și Magdalena Popescu Marin, București, Editura SAECULUM I. O., 2001.
- Chin-Yew Lin, „ROUGE: A Package for Automatic Evaluation of Summaries”, în *Text Summarization Branches Out*, Association for Computational Linguistics, Barcelona, 2004, pp. 74-81.
- Delisle, J., „L'évaluation des traductions par l'historien” în *Meta Journal des traducteurs*, Vol. 46, nr. 2, Les Presses de l'Université de Montréal, Montréal, 2001, pp. 209-226.
- Dicționarul explicativ al limbii române*, Univers Enciclopedic Gold, București, 2016.
- Even-Zohar, I., „Poetics Today. Polysystem Studies”, în *International Journal for Theory and Analysis of Literature and Communication*, Vol. 11, Tel Aviv, 1990.

- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova, „BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”, în *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Vol. 1 (Long and Short Papers), Association for Computational Linguistics, Minneapolis, Minnesota, 2019, pp. 4171-4186.
- Kishore Papineni, Salim Roukos, Todd Ward, Wei-Jing Zhu, „Bleu: a Method for Automatic Evaluation of Machine Translation”, în *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Philadelphia, Pennsylvania, 2002, pp. 311-318.
- Lungu-Badea, G., *Mic dicționar de termeni utilizați în teoria, practica și didactica traducerii*, ed. a III-a, Editura Universitatii de Vest, Timișoara, 2012.
- Denkowski M., Lavie, A., „Meteor Universal: Language Specific Translation Evaluation for Any Target Language”, în *Proceedings of the Ninth Workshop on Statistical Machine Translation*, Association for Computational Linguistics, Baltimore, Maryland, 2014, pp. 376-380.
- Mimansa Jaiswal, Cristian-Paul Bara, Yuanhang Luo, Mihai Burzo, Rada Mihalcea, and Emily Mower Provost, „MuSE: a Multimodal Dataset of Stressed Emotion”, în *Proceedings of the Twelfth Language Resources and Evaluation Conference*, European Language Resources Association, Marseille, 2020, pp. 1499-1510.
- Mounin, G., *Les belles infidèles*, Presses Universitaires de Lille, Lille, 1994.
- Nida, E. A., *Traducerea sensurilor: Traducerea: posibilă și imposibilă*, trad. de R. Dimitriu, Institutul European, Iași, 2004.
- Nida, E. A., Taber, C.R., *The Theory and Practice of Translation*, Vol. VIII, nr. II, E. J. Brill, Leiden, 1982.
- Reiß, K., *Möglichkeiten und Grenzen der Übersetzungskritik*, Max Hueber Verlag, München, 1986.
- Reiß, K., Vermeer, H.J., *Towards a General Theory of Translational Action: Skopos Theory Explained*, trad. de C. Nord, Routledge, London – New York, 2014.
- Schleiermacher, F.D.E., *Des différentes méthodes du traduire* (titlul original: *Ueber die verschiedenen Methoden des Uebersetzens*) trad. de A. Berman, Éditions du Seuil, Paris, 1999.
- Schwerter, S., *D'une culture à l'autre. Le défi de traduire les marqueurs régionaux. TTR: traduction, terminologie, rédaction*, Vol. 29, nr. 1, Association canadienne de traductologie, 2016, pp. 115-138.
- Steiner, G., *După Babel. Aspecte ale limbii și traducerii*, trad. de V. Negoită și Ș. Avădanei, Editura Univers, București, 1983.
- Toury, G., *Descriptive Translation Studies – and beyond*, John Benjamins Publishing Company, Amsterdam-Philadelphia, 1995.

- Varga, D., Halákcsy, P., Kornai, A., Nagy, V., Németh, L., Trón, V., „Parallel corpora for medium density languages”, în *Proceedings of RANLP-05*, Borovets, Bulgaria, 2005, pp. 590-596.
- Venuti, L., *The Translator's Invisibility. A History of Translation*, Vol. II, Routledge, Oxon – New York, 2008.
- Vinay, J.-P., Darbelnet, J., *Stylistique comparée du français et de l'anglais. Méthode de la traduction*, Didier, Paris, 1958.