

LIMITS AND POSSIBILITIES IN AUTOMATIC INTERPRETING**LIMITE ȘI POSIBILITĂȚI ÎN TRADUCEREA AUTOMATĂ****Alexandra ILIE**Școala Doctorală de Studii Filologice,
în cotutelă cu Școala Doctorală în Informatică, Iași
Universitatea „Alexandru Ioan Cuza” din Iași**Abstract**

In the digital era, the communication between the members of the same cultural or social group and between the different cultures represents the main source of evolution in all the domains of activity. With each passing day, it transforms into an instrument for the personal and professional development. Today, the Linguistics faces new challenges imposed by virtual reality. Artificial intelligence, the Natural language processing and Computational linguistics appear among most frequent areas of study which develop a huge interest among the researchers worldwide. As Linguistics and Literature, the Translation Studies domain has also been a mean for the improvement and enrichment of communication skills. The Automatic translation comes to help the traditional one having some possibilities oriented to improve the efficiency and the performance of the human translator work. But, the automatic translation machine help is often limited, because it doesn't manage to overstep the contextual and cultural barriers or to understand the irony or the metaphors of a text.

Résumé

Dans la nouvelle ère digitale la communication entre les membres de la même collectivité culturelle ou sociale et, plutôt, la communication entre les différentes cultures représente le filon principal d'évolution de tous les domaines d'activité, en se transformant d'un jour à l'autre dans un instrument de développement personnel et professionnel. Aujourd'hui, la linguistique se confronte avec des nouveaux défis venus de la part de la réalité virtuelle. L'intelligence artificielle, le traitement automatique du langage naturel et la linguistique computationnelle se retrouvent parmi les domaines d'étude qui développent un intérêt assez grand pour les chercheurs de tout le monde. Comme la littérature et la linguistique, la traduction est considéré un moyen d'amélioration et d'enrichissement des habilités communicationnels. La traduction automatique vient aujourd'hui dans l'appui de la traduction traditionnelle avec des possibilités de plus en plus diverses, orientées vers l'optimisation du travail du traducteur humain. Mais, l'aide venu de la part des moteurs automatiques de traduction a toujours quelques limites, car les barrières imposées par le contexte, la culture, l'ironie, la métaphore, etc. d'un texte ne peuvent pas encore être dépassées.

Rezumat

În era digitală, comunicarea între membrii aceleiași colectivități culturale sau sociale și, mai cu seamă, a comunicării dintre diferitele culturi, reprezintă filonul principal de evoluție al tuturor domeniilor de activitate, transformându-se, de la o zi la alta, tot mai mult, într-un instrument de dezvoltare personală și profesională. Astăzi, lingvistica se confruntă cu noi provocări venite din partea realității virtuale. Inteligența artificială, prelucrarea limbajului natural și lingvistica computațională se regăsesc printre ariile de studiu care dezvoltă un interes tot mai crescut în rândurile cercetătorilor din toată lumea. La fel ca și lingvistica și literatura, traducerea a fost dintotdeauna un mijloc de îmbunătățire și îmbogățire al abilităților de comunicare. Traducerea automată vine astăzi în sprijinul traducerii tradiționale cu posibilități din ce în ce mai diverse, orientate spre eficientizarea muncii unui traducător uman. Însă, ajutorul venit din partea

motoarelor automate de traducere are și unele limite, acestea nereușind, de cele mai multe ori, să depășească barierele contextuale, culturale, ale ironiei, metaforei etc. dintr-un text.

Key words: *automatic translation, decoding, optimization, communication, evolution*

Mots clé: *traduction automatique, décoder, optimisation, communication, évolution*

Cuvinte cheie: *traducere automată, decodare, optimizare, comunicare, evoluție*

Limite și posibilități în traducerea automată

În procesul de cunoaștere a lumii există o diferență de percepere a realității între „cum ne învață ceilalți” să vedem lumea și cum arată realitatea atunci când este „descoperită de noi înșine”. Pentru a căuta singuri sămburele adevărului, trebuie mai întâi să conștientizăm că există unul. Începem să ne punem tot felul de întrebări, întâi nouă, apelând la experiența anterioară proprie, apoi căutăm răspunsuri la ceilalți, apelând la experiența lor anterioară, și abia apoi începe procesul de documentare. Comunicăm pentru a descoperi tot mai mult lumea care ne înconjoară și, pentru a fi mai eficienți, apelăm la experiențele anterioare ale interlocutorului nostru. Atunci când vorbește, omul se raportează la evenimentele din jurul său, cele deja cunoscute de el (devenite amintiri), la cele trăite de ceilalți, dar și la cele viitoare, (în baza unor presupoziii care apar atunci când, în urma unei analize făcute asupra faptelor anterioare, se ajunge la o idee care poate prevesti viitorul, un viitor care a fost probabil „deja trăit” anterior, fie de el, fie de alte generații sau culturi). Punem totul în balanță cu experiențele noastre. Rezultatul acestui proces este „construirea” unei conștiințe unice. Totodată, omul este supus greșelilor din cauza evenimentelor anterioare care îi alterează gândirea, tot la fel cum o și ajută să se dezvolte. Ne întrebăm dacă amintirea evenimentelor anterioare este un proces similar cu două fețe (eroarea și evoluția) sau se împarte în 2 etape diferite? Dacă cea din urmă ipoteză este cea corectă, atunci care sunt elementele care se transformă în erori și care în evoluție sau analiză profundă? În funcție de ce criterii s-ar face această departajare?

Tot la aceleași întrebări, desigur, sub o altă formă, trebuie să „răspundă” și o mașină în procesul de comunicare sau traducere. Majoritatea instrumentelor de traducere automată cunoscute (ex. Google Translate) funcționează pe același principiu: al memoriei și al experiențelor anterioare. În funcție de frecvențele de utilizare ale termenilor și ale participării la actul de traducere al utilizatorilor (interlocutorilor) *Google Translate* își „construiește” o „conștiință” proprie, în baza căreia face anumite alegeri atunci când i se cere să traducă.

Astăzi, nivelul calității traducerilor (mai ales în cazul limbii române) făcute de *Google* nu se ridică la cel al traducerilor efectuate de om (specialist – lingvist – sau nu). Există o diferență foarte mare între traducerile din și în limba română și unele limbi de circulație internațională, cum ar fi engleza sau franceza. Răspunsul la întrebarea „Oare de ce o traducere a aceluiași text sursă (să spunem, din franceză) într-un text țintă în engleză este mai bună decât traducerea lui într-un text țintă în limba română?”, nu este cel la care ne-am așteptat. Nu pentru că engleza se bucură, din ce în ce mai mult, de privilegiul de a fi considerată principala limbă de comunicare în lume (ni se cere un nivel minim de cunoaștere a limbii engleze în toate domeniile chiar și la majoritatea interviurilor de angajare), ci pentru că utilizatorii vorbitori de limbă engleză au îmbunătățit traducerile făcute de motorul de traducere automată, venind cu o revizuire asupra textului primit ori au oferit un *Feedback*. După ce oferă textul în limba țintă, *Google* ne cere părerea și ne pune să alegem dintre următoarele varietate:

„Cum vi se pare?”

- Această traducere este utilă
- Această traducere nu este completă
- Această traducere nu este corectă
- Această traducere nu este utilă”

Mai mult, acesta ne întreabă și dacă avem comentarii sau sugestii în a doua secțiune a paginii de *Feedback*: „Aveți comentarii sau sugestii?”. Odată completat acest câmp cu o traducere corectă, *Google Translate* își va îmbunătăți „experiența de traducere”. Acest efort de corectare este oricum depus de cel/cei care au căutat un mijloc de traducere mai rapid, însă păstrează aceste revizuiți doar în lucrările lor. Mult mai eficient ar fi dacă ar completa această și secțiune. Astfel, timpul de lucru al următorilor utilizatori (printre care se poate regăsi chiar și utilizatorul care a trimis textul ținută corectat), când vor folosi de acest motor de traducere automată pentru traducerea aceluiași text, va fi cu mult scurtat.

Următorul pas în traducerea automată ar fi să nu mai considerăm că întreg traseul de la o limbă A la o limbă B este realizat doar de o mașină, ci ar trebui să învățăm cum putem pune bazele acestei „conștiințe” proprii, cum putem să facem un calculator să „gândească” singur în funcție de experiențele anterioare și să ignore eroarea. Un traducător automat ar trebui să atingă perfecțiunea, iar omul să vină doar cu „suflarea” finală. Dar, ca să putem „spune” unei mașini să funcționeze cu aceleași principii ca ale noastre este nevoie, mai întâi, să înțelegem omul, este nevoie de o comunicare continuă între culturile lumii. Și pentru ca această comunicare să fie posibilă, este inevitabil să nu ne raportăm și la traducere.

Importanța traducerii s-a dovedit în nenumărate domenii de activitate: religie, politică, sociologie, filozofie, știință, literatură, economie etc. Trăim printre traduceri și traducători. A traduce este un serviciu pe care traducătorii îl fac în slujba compatrioților lor. Pentru a se îmbogăți, culturile au nevoie de traducere.

În *A spune cam același lucru* (ECO, 2008, p. 16), Umberto Eco atrage atenția asupra necesității traducerii în zilele noastre. În ultimile decenii s-a înregistrat o creștere continuă a interesului pentru traductologie, globalizarea fiind unul dintre factorii principali care au contribuit la îmbunătățirea calității traducerilor prin plasarea în același mediu a unor grupuri de indivizi, aparținând unor culturi și limbi diferite. Din necesitatea unei mai bune comunicări între aceste grupuri s-a dezvoltat și nevoia de a traduce cât mai mult. Astăzi nu mai susține nimeni că traducerea este imposibilă, așa cum o făcea altă dată Joachim du Bellay sau Henri Meschonnic; traducerea este, în zilele noastre, necesară. Avem nevoie de traducere pentru a fi în pas cu tendințele din orice domeniu de activitate, pentru a putea crește pe scara evoluției. Ea este instrumentul de bază de care ne putem folosi pentru a da jos barierele impuse de diversitățile limbilor.

Astăzi, începe să se impună, tot mai pregnant, un nou domeniu în traducere : Traducerea automată. Chiar dacă poate nu este atât de dezvoltat și în România, la fel ca în marile centre universitare europene, americane și din Asia, subiectul este de o importantă deosebită pentru viitorul traducerii.

Se încearcă o dezvoltare a mașinilor/aplicațiilor de traducere automată. Mecanismul de funcționare este aparent simplu : Aplicația primește un text într-o limbă A și îl va reda, tradus, în limba B cerută, în câteva secunde. Însă, pentru ca acest lucru să fie posibil, este nevoie de foarte multă muncă prealabilă: întocmirea unui corpus și a unei baze terminologice, codarea și decodarea mesajului primit etc. Aceste sisteme de traducere automată se bazează pe inteligența artificială și pe prelucrarea limbajului natural. În general acestea sunt sisteme concepute prin participarea utilizatorilor obișnuiți, fără o pregătire lingvistică. Motiv pentru care și domeniul traducerii automate a devenit, treptat, un loc în care primează cantitatea, ci nu calitatea.

Mesajele nu călătoresc pur și simplu de la emițător la receptor, este nevoie ca ele să fie codate. Traseul urmărit de un mesaj începe cu o primă codare a lui într-o limbă. Apoi, el trebuie să treacă prin canalele de comunicare înainte de a fi primit și decodat de receptor. În toate aceste etape, mesajele pot suferi modificări. Plecând de la transmiterea unui mesaj prin preșele tipografice și ajungând la transmiterea lor folosind telefoanele inteligente, *smartphones*, omenirea a realizat numeroase revoluții în domeniul comunicării.

Astăzi, traducerea începe să aibă un rol foarte important în lingvistică, mai cu seamă în mediile academice. Odată cu creșterea numărului de traduceri și a dezvoltării tehnologiilor s-a produs și o scădere a calității traducerilor și, cu precădere, a traducerilor făcute în mod automat

care, din păcate, de cele mai multe ori, nu pot fi luate nici măcar drept prim suport de lucru. Pentru optimizarea proceselor, considerăm imperativă participarea și conlucrarea specialiștilor din domeniile lingvisticii generale și computaționale, a inteligenței artificiale și prelucrării limbajului natural.

În vederea traducerii unui text este nevoie ca acesta să fie, întâi de toate, înțeles și corect analizat în limba din care provine, limba sursă. Abia după o analiză morfologică, sintactică, lexicală, contextuală etc. făcută în prealabil putem trece la următorii pași. Tipului de text din care face parte textul nostru (găsim analizat în detaliu acest aspect la Katharina Reiss), contextul publicării lui, publicul cărui i se adresează și nivelul lui de educație etc., scopul traducerii (și teoria skoposului la Hans Vermeer), timpul în care a fost scris textul sursă și cel în care se face traducerea, intenția autorului și motivul pentru care a fost scris acel text sunt alte elemente de care trebuie să se țină cont în procesul de traducere. De foarte multe ori, nu mai este necesar să așternem pe foi răspunsul acestor analize. Creierul nostru poate oferi soluțiile într-un timp neașteptat de scurt. Se poate întâmpla ca, din cauza experienței anterioare insuficiente, traducătorul uman să producă, și el, erori. Exemplul cel mai la îndemână fiind analiza de bază a acestor factori de la care a plecat traducerea nuvelei lui Aldous Huxley, *Point Counter Point*. Ca în majoritatea textelor literare, motivul scrierii acestei nuvele nu este unul evident și avem nevoie de investigații amănunțite pentru a-l putea găsi pe cel corect. Tocmai intenția de ironizare a societății nu se regăsea, la momentul respectiv (pe la 1930), printre variantele de interpretare a textului și astfel, nuvela a fost clasată ca făcând parte din categoria operelor superficiale. Interpretarea greșită a unui text sursă duce, inevitabil, la crearea unui text țintă, prin traducere, de o calitate inferioară. Putem acum vedea, mai clar, de ce este nevoie ca un text să fie, mai întâi de toate, înțeles, în întregimea lui, în limba din care provine, folosind mijloacele de analiză specifice limbii respective și abia apoi putem întreprinde traducerea lui.

Prin aceleași procese ar trebui să treacă și un instrument de traducere automată. Deocamdată aceste aplicații folosesc însă, doar algoritmi care se bazează, în cazul celor mai multe dintre ele, pe frecvențele de utilizare, după anumite date statistice. În momentul actual se face apel la traducerea din dicționar a cuvintelor, și nu se ține cont de context și de factorii enumerați mai sus. Pentru a reuși să dezvoltăm procesul acesta sinuos și foarte necalitativ este nevoie să putem face un calculator „să înțeleagă”, în primul rând, gramatica din spatele unui text, morfologia și sintaxa lui. A devenit insuficient să ne concentrăm doar pe vocabular. Dacă vrem să mergem și mai departe, este nevoie să implicăm aspectele legate de context și stilistică. După cum am arătat, poate fi dificil și pentru un om să înțeleagă unele metafore sau ironii fine, ori trimiterea la alte texte și lecturi (tocmai pentru că traducerea depinde de experiența anterioară a unui traducător). Însă, printr-un simplu apel la *BigData*, un motor automat ar putea să extragă, tocmai elementele de care are nevoie pentru o analiză corectă a textului sursă. Minte traducătorului uman se poate adesea înșela și nu poate cuprinde toate informațiile pe care *BigData* le poate conține.

Un bun exemplu (de urmat și de continuat) în acest demers este munca lui Radu Simionescu care a reușit să realizeze, la Iași, un instrument de identificare și separare a unităților de grupuri nominale (Noun Phrase Chunker) în limba română, după ce a fost mai întâi trecut prin filtrul unui PoS Tagger (Part Of Speech tagger) românesc care combină un model de analiză statistic cu unul bazat pe reguli. Dicționarul morfologic a fost extras în mare parte din DexOnline și conține 1,25 milioane de cuvinte distincte. *POS tagger-ul* are o precizie de 96.6% și a fost construit pe baza unei variante corectate a corpusului *1984* de George Orwell. La final, textului de analizat, i s-a aplicat o gramatică *GGs* (Graphical Grammar Studio), orientată spre analiza sintactică. (SIMIONESCU, 2011)

Pentru îmbunătățirea calității traducerilor automate este nevoie de participarea specialiștilor cu o bună pregătire în domeniu și, mai ales, de realizarea unei relații strânse de colaborare între lingvist și informatician. O relație care pare încă greu de stabilit, cu temeri, rețineri și preconcepții din ambele părți. Deși fiecare domeniu, și cel al lingvisticii generale și traducerii tradiționale, și cel al lingvisticii computaționale și traducerii automate prezintă în continuare un interes foarte mare

pentru noul val de cercetători, o colaborare între specialiștii acestor discipline rămâne, poate nu pentru mult timp de acum încolo, un subiect delicat. Un subiect asupra căruia am zăbovit și eu în prezentarea mea. Mesajul pe care îmi doresc să îl transmit odată cu acest articol este că avem nevoie cu toții să stabilim o relație fructuoasă și calitativă între lingvist și informatician, între inginerii de limbaj și inginerii de sisteme automate, o relație care, fără îndoială, va fi de bun augur pentru ambele discipline. Rolul traducerilor automate nu a fost, nu este și nu va fi niciodată acela de a înlocui traducătorul uman. Scopul îmbunătățirii lor este acela de a-l ajuta pe traducător/interpret, de a-i facilita munca, de a eficientiza procesul de traducere în care acesta este angajat. „Suflarea” finală asupra unui text va aparține întotdeauna specialistului uman, riguros pregătit de disciplinele umaniste care deține o formare serioasă în lingvistică generală, lexicologie, morfologie, sintaxă, semantică traductologie, terminologie și nu numai.

Bibliografie

ECO, Umberto, *A spune cam același lucru. Experiențe de traducere*, tradus din italiană de Laszlo Alexandru, Iași, Editura Polirom, 2008

HUXLEY, Aldous, *Point Counter Point*. Garden City, New York, Dalkey Archive Press, 1996

SIMIONESCU, Radu, *Romanian Deep Noun Phrase Chunking Using Graphical Grammar Studio*, 2011, ConsILR conference 2011-2012

