

Silvia-Adriana APOSTOL

Silvia-Adriana APOSTOL

**INTRODUCERE
ÎN
TRADUCEREA ASISTATĂ
DE CALCULATOR**



EDITURA UNIVERSITARIA

Craiova, 2020

Referenți științifici

Prof. univ. dr. Sanda-Marina Bădulescu

Conf. univ. dr. habil. Diana-Adriana Lefter

Copyright © 2020 Editura Universitaria

Toate drepturile sunt rezervate Editurii Universitaria

Descrierea CIP a Bibliotecii Naționale a României

APOSTOL, SILVIA ADRIANA

Introducere în traducerea asistată de calculator / Silvia-Adriana
Apostol. - Craiova : Universitaria, 2020

Conține bibliografie

ISBN 978-606-14-1678-3

81

004

© 2020 by Editura Universitaria

Această carte este protejată prin copyright. Reproducerea integrală sau parțială, multiplicarea prin orice mijloace și sub orice formă, cum ar fi xeroxarea, scanarea, transpunerea în format electronic sau audio, punerea la dispoziția publică, inclusiv prin internet sau prin rețelele de calculatoare, stocarea permanentă sau temporară pe dispozitive sau sisteme cu posibilitatea recuperării informațiilor, cu scop comercial sau gratuit, precum și alte fapte similare săvârșite fără permisiunea scrisă a deținătorului copyrightului reprezintă o încălcare a legislației cu privire la protecția proprietății intelectuale și se pedepsesc penal și/sau civil în conformitate cu legile în vigoare.

Introducere

Atunci când acești doi termeni, **traducere** și **calculator**, apar unul lângă celălalt, publicul larg, care nu posedă cunoștințe de specialitate în domeniul traducerii, își imaginează imediat că este vorba despre vreun program inteligent prin care traducerea este efectuată în mod automat de calculator. Cei care nu lucrează cu limbile străine, în general, și cu traducerea, în mod special, cred că a traduce înseamnă a înlocui cuvintele sau structurile dintr-o limbă cu cuvinte/structuri corespunzătoare din altă limbă. De altfel, acesta a fost și principiul primelor încercări de automatizare a traducerii și tot aceasta a fost și cauza dezamăgirii provocate de faptul că rezultatul primelor încercări de a obține traduceri bune, realizate în mod automat prin programe de traducere automată, nu s-a ridicat la nivelul așteptărilor inițiale. Actul de a traduce este un proces complex care presupune, desigur, o bună decodare (înțelegere) a textului sursă¹ pentru o tot la fel de bună codare (transpunere) a mesajului acestuia în limba țintă pentru a obține un text țintă² care să fie cât mai fidel conținutului și specificului textului original, dar și cât mai natural și, deci, în spiritul limbii țintă. Însă e de necontestat faptul că progresul în domeniul informatic a schimbat într-o mare măsură activitatea traducătorului la toate nivelurile: găsirea de noi clienți, acceptarea cererilor de traducere, comunicarea cu clienții, primirea și transmiterea materialelor, documentarea, cercetarea terminologică, schimbul de baze terminologice/glosare cu alți traducători, înmagazinarea segmentelor traduse în baze de date proprii (memorii de traducere), folosirea memoriilor de traducere puse la dispoziție de clienți sau de alți traducători, comunicarea în cadrul platformelor create pentru comunitatea traducătorilor, pre-editarea și post-editarea textelor, adaptarea aplicată traducerii de softuri sau de jocuri video (localizarea) etc.

¹ Limba sursă = limba în care este redactat textul în forma sa originală.
Text sursă = textul original.

² Limba țintă = limba în care se traduce un text.
Text țintă = textul tradus într-o altă limbă decât cea în care a fost scris inițial.

Prin urmare, dacă traducerea automată de foarte bună calitate care să nu necesite deloc intervenția umană nu este încă posibilă, calculatorul oferă traducătorului instrumente care îi pot ușura munca, ajutându-l să fie mai eficient, să traducă bine, mai mult și în timp mai scurt. Aceasta ar putea fi, într-o viziune de ansamblu, definiția largă a **traducerii asistate de calculator (TAC)**, obiectul lucrării de față, care își propune să adauge abordării teoretice a conceptelor cu care operează TAC prezentări practice și să creeze astfel un continuum practico-teoretic și teoretico-practic.

Lucrarea se adresează studenților și tuturor celor interesați de domeniul traducerii și de schimbările pe care dezvoltarea tehnologiei le-a adus și le-a impus în practica traducerii.

Prima dovadă clară a faptului că există o legătură între practica traducerii și domeniul informatic este aceea că nimeni nu-și imaginează că o traducere (chiar și cea literară) ar mai putea fi predată scrisă de mână.

Evident, traducerea asistată de calculator nu înseamnă (doar) asta și nu toate domeniile de specialitate se pretează la folosirea aplicațiilor informatice de asistare a traducătorului. Traducerea literară rămâne, în ansamblul ei, un domeniu rezistent la astfel de sisteme informatice. Cu toate acestea, chiar și pentru traducerea literară, rezultatele unui concordanțier, de exemplu, pot fi revelatoare pentru importanța folosirii unui anumit cuvânt în cadrul operei literare supuse decodării și (re)codării interlingvistice prin indicarea tuturor ocurențelor aceluși cuvânt și a contextelor în care acesta apare.

Multe dintre instrumentele de investigare lingvistică folosite în lingvistica computațională pentru activități legate de cercetarea limbajului uman sunt extrem de utile în procesul de traducere a textelor de specialitate. În prezent, este aproape de neimaginat, de exemplu, ca un traducător profesionist să nu recurgă la extragerea termenilor dintr-un corpus pentru pregătirea procesului de traducere (extragerea bilingvă a termenilor, crearea de glosare) sau la sistemele bazate pe memorii de traducere pentru editarea traducerii propriu-zise.

Considerăm că, la început de drum, orice traducător are nevoie de un plan de călătorie în aventura traducerii cu ajutorul calculatorului: mai exact, ar trebui să știe la ce să se aștepte (ce înseamnă TAC), ce trebuie să pună în bagaje (cunoștințe), ce mijloace de transport alege (diferite instrumente) și care sunt regulile de utilizare a acestora, care sunt plusurile și minusurile acestor mijloace, costuri, variante *low* sau *no-cost*, disponibilitate etc.

Pe tot parcursul lucrării abordarea teoretică este dublată de exemple practice, tot așa cum activitățile practice sunt însoțite de explicații teoretice.

Capitolul I

Ce este traducerea asistată de calculator?

1.1. Traducerea asistată de calculator (TAC) și Traducerea automată (TA)

1.1.1. TAC și TA conform *Dicționarului terminologic* al Oficiului de limbă franceză din Quebec

Vom porni în definirea conceptului de **traducere asistată de calculator** (TAC) de la definiția din dicționarul terminologic (*Le grand dictionnaire terminologique*) realizat de Oficiul de limbă franceză din Quebec:

*Traduction assistée par ordinateur: Technique permettant d'engendrer des textes en langue cible à partir d'une langue source par un système informatique avec l'assistance d'un traducteur humain.*¹

Plecând de la această definiție, înțelegem că TAC ar fi o tehnică de traducere „prin care sunt generate texte într-o limbă țintă plecând de la o limbă sursă de către un sistem informatic, asistat de un traducător uman”. Cu alte cuvinte, sistemul informatic ar fi cel care traduce textele în limba țintă cu ajutorul traducătorului uman.

Fișa terminologică a termenului TAO (*Traduction assistée par ordinateur*) prezentată în dicționarul terminologic amintit mai sus precizează că asistența umană poate interveni în trei etape ale procesului de traducere: în

¹ Office québécois de la langue française, *Le Grand dictionnaire terminologique*, 1997, disponibil online http://www.granddictionnaire.com/ficheOqlf.aspx?Id_Fiche=8394081 ; ultima accesare 08/10/2018.

Pentru această parte a definițiilor TA și TAC, am ales să păstrăm și să cităm definițiile în limba originală, întrucât comentariile pe care le facem pe baza acestora implică și traducerea lor în limba română. Acolo unde comentariile reiau doar anumite aspecte ale definiției, traducerea completă este dată în notă de subsol.

etapa de pre-editare (adnotarea sau transformarea textului sursă), în etapa de traducere propriu-zisă (sistem interactiv) și în etapa de post-editare (revizuirea traducerii).¹

Același dicționar terminologic prezintă următoarea definiție a **traducerii automate**:

*Technique informatique permettant d'obtenir de façon automatique, sans l'intervention d'un traducteur humain, une traduction de textes d'une langue source vers une autre langue.*²

Conform acestei definiții, traducerea automată este „o tehnică informatică ce permite obținerea unei traduceri în mod automat, fără intervenția traducătorului uman”. Fișa terminologică a termenului TA (*traduction automatique*) precizează în notă că traducerea rezultată în mod automat nu este decât o traducere aproximativă a textului original, la care putem apela doar pentru a avea o viziune de ansamblu asupra conținutului textului. Exemplul dat este acela al folosirii traducerii automate pentru o înțelegere globală, brută, a subiectului unei pagini Web într-o limbă pe care nu o cunoaștem deloc³.

Observații:

Reținem că cele două definiții de mai sus prezintă traducerea asistată de calculator și traducerea automată ca două modalități de abordare a activității de traducere, două tehnici informatice folosite pentru obținerea unui text tradus dintr-o limbă sursă într-o limbă țintă, prima (TAC), cu asistența traducătorului uman, cealaltă (TA), fără intervenția traducătorului uman.

¹ *Idem.*

² *Ibidem*, http://www.granddictionnaire.com/ficheOqlf.aspx?Id_Fiche=8395113; ultima accesare 08/10/2018.

³ Office québécois de la langue française, *Le Grand dictionnaire terminologique*, 1997, disponibil online http://www.granddictionnaire.com/ficheOqlf.aspx?Id_Fiche=8395113; ultima accesare 08/10/2018.

Insistența asupra componentei informatice în definiția TAC din *Grand dictionnaire terminologique*, în care traducătorul este cel care asistă calculatorul (programul informatic) și intervine asupra procesului de traducere la diferite niveluri, și nu invers, trimite mai degrabă la conceptul de « traduction machine assistée par l'homme » (*Human-assisted machine translation*; traducere automată asistată de utilizator). Reținem, în schimb, lipsa intervenției utilizatorului (traducătorului uman) în definiția TA, care ar corespunde conceptului de *Fully automatic machine translation* (traducere automată completă). Până nu demult, aplicațiile de TA erau folosite doar pentru înțelegerea globală¹ a temei unui text. Deși dicționarul terminologic amintit aici vorbește despre utilizarea limitată a acestui tip de tehnologie, observație pe care o regăsim și în alte lucrări², progresele continue înregistrate de sistemele de traducere automată, ca urmare a interesului sporit pentru dezvoltarea acestor aplicații din partea instituțiilor internaționale și a marilor companii, impun o reconsiderare a rolului traducerii automate în industria limbilor și a traducerilor³. De altfel, rolul tot mai important al traducerii automate transpare și din faptul că sistemele de traducere asistată de calculator includ acum și funcția de traducere automată, la care traducătorul uman poate apela pentru o primă schițare a traducerii, a cărei calitate depinde de motorul de TA utilizat și de perechile de limbi aflate în contact.

Observăm totuși că, în general, atunci când companiile sau instituțiile folosesc traducerea automată pentru versiunile în alte limbi ale siturilor proprii sau pentru alte documente, își iau măsuri de precauție și precizează sau, mai bine spus, avertizează asupra faptului că s-a folosit doar traducerea automată.

¹ « des aides à la compréhension », după cum le numește Bădulescu.

Bădulescu, Sanda-Marina, *Traduction assistée par ordinateur pour les langues de spécialité*, București, Editura Bren, 2004, p.16.

² Vezi Gasperi, F., cap. „Introduzione alla traduzione automatica”, în BERSANI BERSELLI, Gabriele (dir.), *Usare la traduzione automatica*, CLUEB Editore, 2011, pp. 13-32.

³ Vezi rezultatele studiilor realizate în 2019 și în 2020 de către EUATC (European Union of Associations of Translation Companies – Uniunea Europeană a Asociațiilor de Firme de Traducere), prezentate în Anexa 1.

1.1.2. TAC și TA conform definiției lui Lynne Bowker

Lynne Bowker face următoarea distincție între **traducerea automată** („human-assisted machine translation - HAMT” sau, mai scurt, „machine translation” – MT) și **traducerea asistată de calculator** [„computer-aided (or-assisted) translation” - CAT, cunoscută și sub numele de „machine-assisted (or-aided) human translation” – MAHT]:

The major distinction between MT and CAT lies with who is primarily responsible for the actual task of translation. In MT, the computer translates the text, though the machine output may later be edited by a human translator. In CAT, human translators are responsible for doing the translation, but they may make use of a variety of computerized tools to help them complete this task and increase their productivity. Therefore, whereas MT systems try to replace translators, CAT tools support translators by helping them to work more efficiently.¹

Distincția pe care Lynne Bowker o face între traducerea automată și traducerea asistată de calculator se bazează pe rolul pe care îl joacă traducătorul uman, respectiv calculatorul, în procesul de traducere. Astfel, în **traducerea automată**, cel care realizează traducerea este calculatorul. Traducătorul uman poate interveni ulterior pentru a corecta rezultatul obținut prin traducerea automată. De cealaltă parte, în **traducerea asistată de calculator**, cel care realizează traducerea este traducătorul uman, care

¹Cui îi revine sarcina de a efectua traducerea propriu-zisă: iată în ce constă principala deosebire dintre TA și TAC. În TA, calculatorul este cel ce traduce textul, rezultatul astfel obținut putând fi editat ulterior de un traducător uman. În TAC, cel care efectuează traducerea propriu-zisă este traducătorul uman, însă el poate utiliza în acest scop o serie de instrumente informatice, cu ajutorul cărora devine totodată mai productiv. Așadar, dacă sistemele de TA încearcă să-l înlocuiască pe traducător, instrumentele de TAC vin, în schimb, în sprijinul acestuia, ajutându-l să lucreze mai eficient. (t.n.)

Bowker, L., *Computer-aided Translation Technology: A Practical Introduction*, Didactics of Translation Studies, University of Ottawa Press, Ottawa, 2002, p. 4.

poate folosi în acest scop diverse instrumente. De asemenea, Lynne Bowker subliniază și creșterea productivității ca urmare a utilizării instrumentelor de TAC.

1.1.3. Concluzii

Etapa de „pesimism excesiv”¹ care a marcat cercetarea în domeniul programelor informatice de automatizare a traducerii după așa-zisul eșec al proiectului ambițios al lui Warren Weaver de a obține o traducere automată completă de înaltă calitate (FAHQT – *Fully Automatic High Quality Translation*) a fost urmată de o abordare mai realistă, prin care s-au căutat soluții informatice practice, care să răspundă efectiv nevoilor traducătorului². Este vorba de crearea de sisteme integrate (sau posturi de lucru integrate) care să-l ajute pe traducător să efectueze traduceri de înaltă calitate în timp scurt, deci cu eficiență maximă.

Definiția largă a TAC ar include toate instrumentele informatice pe care traducătorii le pot folosi în activitatea de traducere: chiar și simpla folosire a calculatorului pentru editarea traducerii ar intra în această accepție.

Într-o accepție mai restrânsă, TAC se referă la aplicații care au ca principiu de bază **memoria de traducere**. Instrumentele TAC au la bază principiul memoriilor de traducere, însă nu se reduc la asta, ele sunt de fapt aplicații care permit o comandă centralizată a unei serii întregi de instrumente: crearea și gestionarea memoriilor de traducere, aliniere de texte, gestiune terminologică (crearea de baze de date terminologice, extragerea termenilor dintr-un document, importarea și exportarea bazelor de date, acces direct la baze de date terminologice integrate în aplicație), editare de text, post-editarea

¹ Ladmiral, J.-R., « Le traducteur et l'ordinateur », in *Langages*, 28^e année, no. 116, 1994. Le traducteur et l'ordinateur, pp. 5-19 ; doi :<https://doi.org/10.3406/lgge.1994.1691>, https://www.persee.fr/doc/lgge_0458-726x_1994_num_28_116_1691, ultima accesare 09/10/2018.

² Pentru un scurt istoric al traducerii automate și al evoluției traducerii asistate de calculator, vezi Bădulescu, S.-M., « Historique de la TAO », în *op.cit.*, pp. 14-18.

traducerii automate integrate în aplicație, corectare ortografică, editare de text predictiv, gestiunea proiectelor de traducere etc.

1.2. Concepte-cheie în TAC

1.2.1. Memorie de traducere

Memoria de traducere este o bază de date (o arhivă) în care sunt salvate segmente de text în limba sursă și segmentele de text corespunzătoare traduse în limba țintă. Memoria de traducere poate fi creată de traducătorul uman sau poate fi generată prin sisteme de aliniere a textelor.

În cazul memoriei de traducere create de traducătorul uman, înainte de a începe să traducă un text, traducătorul creează o memorie de traducere goală pe care o va alimenta pe parcurs. Practic, aplicația de memorie de traducere integrată în instrumentul TAC împarte textul în segmente de text (delimitate la nivelul propoziției sau al frazei/al paragrafului, dar cu posibilitatea de a le micșora sau lărgi). De fiecare dată când un segment de text este tradus, acesta este asociat segmentului corespunzător tradus în limba țintă și împreună formează o unitate de traducere¹ (prescurtată și *TU* din corespondentul în limba engleză *translation unit*). Unitățile de traducere astfel formate sunt salvate și înmagazinate în memoria de traducere.

În timpul procesului de traducere, memoria este consultată în mod automat pentru a verifica dacă pentru segmentul în curs de traducere există deja un segment identic sau similar în memoria de traducere respectivă. Utilizatorul (traducătorul) are posibilitatea de a accepta propunerea de

¹ Într-o accepție traductologică mai largă, unitatea de traducere este, după cum o definește Teodora Cristea, „o entitate de sens cu față dublă”, întrucât presupune interpretarea intralinguală în limba sursă, pe de-o parte, și interpretarea intralinguală în limba țintă, pe de altă parte, în căutarea unui raport de echivalență interlinguală”:

« L’UT est une entité à double face, car les deux codes, celui de la langue source et celui de la langue cible sont réunis par une relation interlinguale qui émerge à la suite de deux exégèses intralinguales à travers lesquelles on recherche la signification commune »

Cristea, T., *Stratégies de la traduction*, Editura Fundației „România de mâine”, București, 2000, p. 19.