

Pachetul de programe *Trados*

Sanda Cherata
Universitatea Babeş-Bolyai Cluj-Napoca
Facultatea de Litere

Introducere

În procesul deosebit de accelerat de dezvoltare pe toate planurile, la care asistăm în momentul de față, echipele de documentare, de traducere și de localizare și din cadrul întreprinderilor sunt supuse unor cereri de performanțe din ce în ce mai exigente. Așa cum firmele producătoare fac eforturi să răspundă cerințelor de competitivitate pe piața internațională, aceste echipe trebuie să țină pasul în același ritm pentru a furniza, cu aceeași eficiență și la un înalt nivel profesional, localizările și traducerile documentelor. Fără instrumente informatice care să-i asiste pe tot parcursul activității lor, traducătorii nu mai pot face față acestor exigențe.

Aplicațiile informatice de asistare a traducătorilor¹ existente în momentul de față sunt foarte variate: unele sunt dedicate rezolvării unei singure sarcini (memorie de traducere, bază de date terminologice etc.), altele integrează într-un singur pachet mai multe instrumente care concluează pentru atingerea scopului de asistare cât mai eficientă a traducătorilor.

Pachetul **Trados** are avantajul că integrează o sumă de instrumente utile de asistență a traducătorilor: memoria de traducere (**Translator's WorkBench**), un sistem de editare pentru textele cu marcaje (**TagEditor**), sistemul de baze de date terminologice (**MultiTerm**), un instrument de extragere a terminologiei (**ExtraTerm**), un aliniator de texte (**WinAlign**) precum și instrumente de gestiune a proiectelor de traducere.

1 Gestiunea memoriilor de traducere (Translator's TM WorkBench)

O memorie de traducere este, de fapt, o bază de date în care se memorează segmentele de text în limba sursă împreună cu segmentele de text corespunzătoare traduse în limba țintă. Dacă, la traducerea unui document, se folosește o memorie de traducere creată anterior, pentru toate segmentele de text care se regăsesc în memoria de traducere, traducerea este furnizată automat utilizatorului. În felul acesta, frazeologia și terminologia unui domeniu sunt tratate consecvent și unitar în toate documentele care tratează domeniul respectiv.

În **Trados** se pot trata texte scrise în orice limbi al căror set de caractere e suportat de sistemul de operare și în cele mai utilizate formate: documente Word, documente în format text îmbogățit (RTF), texte în format ML (cu marcaje: HTML, SGML, XML etc.). Asupra specificității tratării textelor cu marcaje vom reveni.

Pentru utilizator, gestiunea memoriei de traducere se face în mod similar, indiferent de formatul textului tratat.

În timpul traducerii, **WorkBench** divide textul sursă în **segmente**. Acestea sunt, de obicei, delimitate de terminatorii de propoziție. Utilizatorul poate lărgi sau restrânge după dorință segmentul creat de **WorkBench**. Trebuie să menționăm aici un avantaj care nu e de neglijat: opțiunile de formatare și de punere în pagină a segmentului din textul sursă sunt preluate automat pentru textul tradus în limba țintă.

¹ A se vedea Cherata (2004)

Unul câte unul, segmentele sunt afișate într-un câmp evidențiat. Imediat sub câmpul segmentului în limba sursă apare un câmp evidențiat într-o culoare diferită în care utilizatorul introduce traducerea corespunzătoare a segmentului. Fiecare segment de text din limba sursă este legat de traducerea lui în limba țintă și formează împreună cu ea o **unitate de traducere** [en: *translation unit*, uneori prescurtat prin *TU*], care se înregistrează în memoria de traducere.

La tratarea unui segment al textului sursă, acesta este comparat cu segmentele existente deja în memoria de traducere. Dacă același segment, sau un segment cu conținut asemănător, este găsit în memorie, traducerea corespunzătoare, din cadrul unității de traducere găsite, este propusă utilizatorului. Acesta o poate accepta sau modifica, după caz. Această reutilizare a fragmentelor de text deja traduse crește substanțial viteza, consistența și eficiența procesului de traducere.

1.1 Crearea memoriilor de traducere

Când nu există o memorie de traducere (fie din cauză că se traduce un document cu un conținut diferit față de cele tratate până în momentul respectiv, fie pentru că până în acel moment nu a fost utilizat pachetul **Trados**), se cere **crearea** unei memorii **noi**. Aceasta este inițial vidă (nu conține nici o unitate de traducere) și se îmbogățește pe măsura traducerii documentului, prin procedeul arătat mai sus.

Pe lângă această situație tipică, pachetul **Trados** oferă instrumentele și mijloacele prin care se pot crea memorii de traducere plecând de la perechi de documente în limba sursă și traducerile corespunzătoare.

De exemplu, dacă există un număr de documente deja traduse fără mijloace de asistare a traducătorului, programul **WinAlign** (furnizat ca instrument independent în cadrul pachetului **Trados**) realizează automat (cu posibilitatea intervenției utilizatorului) legăturile dintre segmentele de text din limba sursă și segmentele corespunzătoare traduse în limba țintă (operație numită **aliniera** textelor, de unde și denumirea programului). O funcție a programului **WinAlign** (meniul **File**, comanda **Export Project**) permite salvarea, în format text, a rezultatului alinierii segmentelor. Din **WorkBench**, acesta poate fi importat (meniul **File**, comanda **Import**) în memoria de traducere deschisă (care poate fi una nouă sau una existentă).

1.2 Utilizarea și întreținerea memoriilor de traducere

O memorie de traducere poate fi folosită cu eficiență la traducerea documentelor cu conținut înrudit cu al textelor din care aceasta a fost creată. În acest scop este suficient ca, înainte de a se deschide un document spre a fi tradus, să se lanseze programul **WorkBench** și să se deschidă memoria de traducere adecvată. Pentru orice segment găsit în memoria de traducere se furnizează traducerea corespunzătoare. Gradul de asemănare a segmentului de text din documentul de tradus cu cel găsit în memoria de traducere poate varia de la corespondența 100% – cel mai înalt – până la un grad de corespondență specificat în procente de utilizator [en: *fuzzy match*]. Utilizatorul poate prelua traducerea furnizată și, dacă e cazul, o poate modifica. Dacă segmentul nu se identifică 100% cu cel existent în memorie, este memorat împreună cu traducerea lui. De asemenea, dacă se modifică traducerea unui segment din memorie, noua traducere este și ea înregistrată. Utilizatorul trebuie să specifice, la crearea memoriei de traducere, dacă se permit sau nu traduceri multiple ale aceluiași segment de text sursă. În funcție de această opțiune, dacă se dă o nouă traducere a unui segment deja existent în memorie se păstrează numai ultima traducere sau se memorează noua traducere într-o unitate de traducere nouă.

O facilitare de mare utilitate pentru traducători este aceea de inserare în documentul tradus a elementelor din text care fie sunt specifice localizării, fie, în general, trebuie copiate

ca atare fără traducere (în engleză se numesc *„placeables”*). Acestea sunt mărimi numerice (care, eventual, ar trebui localizate, în funcție de unitățile de măsură), date calendaristice, imagini (acestea trebuie preluate ca atare, fără traducere), hiper-legături (care, eventual, pot fi editate dacă fac trimitere la un punct al documentului) etc. Astfel de elemente sunt identificate de **WorkBench** și pot fi inserate automat în document la poziția dorită.

1.2.1 Analiza documentelor și pre-traducerea automată

O funcție foarte utilă a programului **WorkBench** este aceea de analiză a documentelor (**Analyze**) care permite utilizatorului să ia decizia, înainte de a începe traducerea unuia sau mai multor documente, dacă este eficient sau nu să utilizeze o memorie de traducere existentă. Mai mult decât atât, traducătorul poate aprecia mai exact cantitatea de muncă pe care o are de depus efectiv în procesul de traducere (pentru a stabili costul traducerii) și poate efectua o pre-traducere automată a textului.

Astfel, prin funcția de **analiză a documentelor**, utilizatorul poate afla câte segmente din documentul sau documentele analizate se găsesc deja în memoria de traducere (împreună, evident, cu traducerile corespunzătoare) precum și gradul de corespondență dintre acestea (câte corespund 100%, câte au o corespondență parțială și în ce grad, câte segmente nu au corespondent). De asemenea, este indicat numărul de repetiții ale segmentelor (și cuvintelor) care nu se regăsesc în memoria de traducere; un grad înalt de repetiții este un indiciu al eficienței utilizării memoriei de traducere.

Dacă analiza documentului arată un număr mare de segmente găsite ca potrivindu-se 100% cu cele din memoria de traducere, utilizatorul poate cere o pre-traducere a documentului, în cursul căreia toate segmentele în cauză vor fi traduse automat, fără a fi necesară deschiderea documentului pentru traducerea interactivă. Pentru aceasta, utilizatorul dă comanda **Translate** din meniul **Tools**, în care poate specifica documentele care să fie traduse automat și gradul de corespondență dintre segmentele din documentele de tradus și cele existente în memoria de traducere. Pentru traducerea completă a documentului se procedează în mod normal: după deschiderea lui, se poate da comanda **Open Next no 100%/Get** pentru a se deschide primul segment pentru care nu există traducere.

1.2.2 Întreținerea și reorganizarea memoriei de traducere

Este util ca, din când în când, să se analizeze și, eventual, să se actualizeze conținutul memoriei de traducere. Prin actualizare se înțelege editarea (modificarea conținutului) sau chiar ștergerea unor segmente din memorie.

Prin comanda **Maintenance** din meniul **File** se poate vedea conținutul tuturor unităților de traducere (segmentele cu traducerile lor). Pentru a edita sau șterge o unitate de traducere se dă comanda corespunzătoare.

O funcție foarte utilă a comenzii de întreținere este localizarea și înlocuirea automată a traducerii unui cuvânt sau expresii cu o altă traducere. O astfel de necesitate poate apărea în urma revizuirii traducerii.

Periodic, de obicei după adăugarea a 500-1000 de unități de traducere, este bine ca memoria de traducere să fie reorganizată. Această operație realizează rearanjarea unităților de traducere și „curățirea” memoriei de „balast”. Această operație este necesară mai degrabă pentru „bucătăria internă” a memoriei de traducere, dar utilizatorul poate simți efectul ei în promptitudinea cu care funcționează programul **WorkBench**.

1.3 Gestiunea terminologiei în procesul de traducere

Pe lângă facilitatea utilizării fragmentelor de text traduse, în **WorkBench** există posibilitatea de a gestiona terminologia într-un mod cât se poate de eficient.

Astfel, o bază de date terminologice **MultiTerm** poate fi deschisă și rămâne activă pe parcursul procesului de traducere. Operația de deschidere a bazei se face din meniul **Options**, comanda **Term Recognition Options**. Din acest moment, funcția de recunoaștere terminologică este activă, iar termenii vor fi căutați în baza de date specificată.

Dacă în baza terminologică există înregistrat un termen care apare în segmentul în curs de traducere, acesta este semnalat ca atare și, în fereastra corespunzătoare afișării datelor terminologice este vizualizat termenul respectiv împreună cu echivalenții lui în limba țintă. Utilizatorul poate cere să vadă conținutul întregii fișe terminologice în cauză. Traducerea termenului în limba țintă se poate insera imediat, la poziția cursorului în textul țintă. Mai mult decât atât, între **WorkBench** și **MultiTerm** există posibilitatea de comunicare foarte comodă. Dacă utilizatorul identifică în textul de tradus un termen care nu există în baza de date terminologice, el poate cere înscrierea lui în bază, împreună cu corespondentul său în limba țintă. În felul acesta se poate actualiza interactiv conținutul bazei terminologice, pe parcursul procesului de traducere.

Datorită integrării celor două instrumente (memoria de traducere și baza de date terminologice), terminologia este utilizată în mod consecvent, iar îmbogățirea bazei terminologice se face interactiv, fără a părăsi activitatea de traducere.

1.4 Livrarea documentului tradus

Procesul de traducere se desfășoară în mai multe etape, pe parcursul cărora se operează modificări ale traducerii efectuate. Dacă modificările sunt efectuate în segmentele deschise cu **WorkBench**, acestea sunt reținute în memoria de traducere. În caz contrar (dacă utilizatorul poziționează cursorul la poziția dorită și introduce modificările, fără a deschide segmentul) acestea nu sunt adăugate memoriei de traducere.

Pe tot parcursul procesului de traducere asistată de **WorkBench**, în document sunt păstrate atât segmentele de text sursă cât și segmentele de text traduse în limba țintă; mai mult, în textul documentului sunt introduse marcaje delimitatoare de segment. Păstrarea documentului în forma bilingvă este necesară pe tot parcursul traducerii, deoarece în acest fel **WorkBench**-ul identifică segmentele și unitățile de traducere.

Când traducerea documentului este validată și acesta este gata de a fi predat clientului, documentul trebuie „curățat” de toate elementele introduse de **WorkBench** (inclusiv de textul în limba sursă). Acest lucru se face prin operația **Clean Up**, care trebuie să fie ultima activitate înainte de livrarea traducerii. Pentru aceasta, documentul trebuie salvat (e bine ca salvarea să fie făcută sub un alt nume, pentru a avea și versiunea bilingvă, în caz că se va reveni la procesul de traducere) și închis. Se deschide apoi memoria de traducere utilizată pentru documentul respectiv și, din meniul **Tools** se dă comanda **Clean Up**. În fereastra corespunzătoare, se specifică documentul sau documentele care trebuie curățate și opțiunile de actualizare, după care se dă comanda **Clean Up**. Utilizatorul trebuie să fie atent la specificarea opțiunilor:

- **Don't Update** – modificările care au fost făcute în documentul tradus fără controlul programului **WorkBench** nu vor fi reținute în memoria de traducere;
- **Update TM** – modificările din cadrul unităților de traducere vor fi înregistrate în memoria de traducere;
- **Update Document** – modificările efectuate numai în textul tradus (fără asistența **WorkBench**-ului) vor fi anulate și vor fi înlocuite cu traducerile existente în memoria de traducere;

- **Dont't Clean Up** – unitățile de traducere pentru care segmentele de text țintă nu coincid cu cele corespunzătoare lor înregistrate în memoria de traducere vor rămâne nemodificate și nu sunt îndepărtate elementele introduse de **WorkBench**.

2 Editarea textelor în formate cu marcaje - TagEditor

Textele elaborate într-un format cu marcaje [en: *tags*] – HTML, SGML, XML etc. necesită o tratare specială în timpul editării. În pachetul **Trados** există programul **TagEditor** creat în acest scop.

TagEditor creează mediul de traducere pentru fișiere *Microsoft Excel* și *PowerPoint*, pentru documentele în format cu marcaje – *HTML*, *SGML*, *XML* – ca și pentru documente *Tagged RTF* și *TradosTag*.

Aceste tipuri de fișiere conțin, pe lângă textul propriu-zis, care trebuie tradus, și informații meta-textuale – sub formă de marcaje – care, în general, nu trebuie traduse.

TagEditor-ul convertește automat aceste fișiere într-un format numit **TRADOStag**, în care informațiile de aspect și de formatare sunt separate de cele care constituie textul de tradus.

Interfața programului **TagEditor** este cât se poate de convenabilă pentru acest gen de texte: diversele tipuri de marcaje sunt prezentate diferit, astfel încât utilizatorul poate le distinge ușor pe cele al căror conținut trebuie tradus de cele care trebuie să rămână neschimbate.

Fiecare tip de fișier are marcajele lui proprii.

Traducerea fișierelor editate cu **TagEditor**-ul se face în același mod ca și a documentelor Word. Se lansează **Translator's WorkBench** și se creează o memorie nouă sau se deschide una existentă. Dacă e necesar, se stabilesc opțiunile pentru gestiunea terminologiei. Se lansează în execuție programul **TagEditor** și se deschide documentul de tradus. În principiu, procesul de traducere se desfășoară ca și pentru documentele Word și sunt accesibile toate funcțiile de gestiune a memoriei și de gestiune terminologică.

În plus, **TagEditor** are funcții specifice pentru gestionarea marcajelor. Ele sunt puse în evidență prin casete grafice specifice, care disting marcajul de deschidere de marcajul corespunzător de închidere.

Marcajele externe reprezintă, de obicei, elementele structurale ale documentului (stilurile paragrafelor, elemente ancorate în document: tabele, frame-uri etc.). De obicei, astfel de elemente nu sunt mutate și rareori se adaugă sau se șterg în documentul tradus. Ele apar, în mod obișnuit, la începutul liniei. **TagEditor**-ul le evidențiază în culoare gri cu chenar negru și le ignoră pe parcursul traducerii.

Marcajele interne reprezintă indicații de formatare a textului (font, caractere speciale etc.). Majoritatea marcajelor interne pot fi mutate pentru a modifica formatarea textului sau pot fi adăugate sau șterse în timpul traducerii. De obicei, ele apar în interiorul paragrafelor de text care trebuie traduse. **TagEditor**-ul le evidențiază în culoare gri cu chenar roșu. Deși ele însele nu trebuie traduse, ele trebuie inserate în textul tradus în poziția corectă pentru a corespunde necesităților traducerii.

Modul de vizualizare a conținutului marcajelor poate fi stabilit de utilizator: afișarea simplă a marcajului fără conținut, afișarea parțială a conținutului sau afișarea conținutului complet al marcajului.

La traducerea documentelor în formate cu marcaje este esențial ca textul tradus să conțină aceleași marcaje ca textul sursă. Mai mult, dacă accidental se șterge un marcaj de deschidere sau de închidere, fără a se șterge marcajul-pereche, întreaga structură a documentului poate fi compromisă. **TagEditor**-ul are două facilități foarte utile pentru astfel de situații: utilizatorul poate cere protecția marcajelor (**Tag Protection**) și, la sfârșit, poate face o verificare a consistenței marcajelor (**Tag Validation**).

De asemenea, dacă se dorește introducerea unor marcaje sau elemente noi în textul tradus, acest lucru se poate face simplu, deoarece în meniul **Edit** există comanda **Insert**

Tag, respectiv **Insert Entity**, de unde se poate alege marcajul sau entitatea care trebuie introduse. Pentru ca această operație să fie posibilă este necesar să se dezactiveze protecția marcajelor. După inserarea marcajelor dorite, este bine să se reactiveze protecția marcajelor.

O altă facilitate care ajută mult utilizatorul este cea care permite vizualizarea textului sursă și, respectiv, a celui tradus în modul în care se văd ele în realitate. Astfel, din bara cu unelte de vizualizare, afișată în partea de jos a ferestrei **TagEditor**-ului, se poate cere afișarea textului sursă, sau atât a textului sursă cât și a celui tradus, sau numai a textului tradus.

În plus, **TRADOSTag Viewer** este un program care permite utilizatorului să previzualizeze sau să imprime din **Internet Explorer** orice document **TRADOSTag** (cu extensia .ttx).

Pentru salvarea fișierelor în **TagEditor** se dă comanda **Save Bilingual As** din meniul **File**. Fișierul bilingv este salvat în format **TRADOSTag** (.ttx) și este analog fișierului Word salvat fără operațiunea de curățare.

La sfârșitul procesului de traducere fișierele bilingve (.ttx) trebuie **curățate**, pentru a se păstra numai forma finală a textului tradus.

Dacă toate modificările aduse traducerii s-au efectuat în **TagEditor** prin operații de deschidere/închidere de segment, atunci este suficient ca din meniul **File** să se dea comanda **Save Target As**, care are efectul operației **Clean Up** pentru documente Word.

Dacă însă modificările traducerii s-au făcut în textul țintă fără deschidere/închidere de segment, atunci ele nu au fost reținute în memoria de traducere. În acest caz, operația de curățare a textului se face în același mod ca și pentru documentele Word (operația **Clean Up**).

3 Sistemul de gestiune a bazelor de date terminologice **MultiTerm**

Aplicația **MultiTerm** este ea însăși un sistem complex, destinat creării, dezvoltării și interogării bazelor de date terminologice.

MultiTerm poate fi folosit de sine stătător, pentru crearea, actualizarea și consultarea bazelor de date terminologice.

Pe de altă parte, **MultiTerm** poate fi integrat în editorul de texte Word (fără a se utiliza memoria de traducere prin **WorkBench**). În acest caz, se poate deschide o bază de date terminologice (prin apăsarea unui buton din bara cu unelte **MultiTerm**); pe parcursul editării documentului, baza terminologică poate fi consultată (se poate căuta o anumită intrare), poate fi afișată întreaga fișă terminologică și termenul corespondent în limba țintă poate fi inserat automat în document la poziția cursorului.

Cel mai eficient mod de integrare a unei baze de date **MultiTerm** este cel oferit de **WorkBench** (prezentat în paragraful 1.3) asigurându-se astfel funcția de gestiune terminologică în cazul utilizării memoriei de traducere.

Aplicația **MultiTerm** fiind foarte complexă, ne vom limita la prezentarea celor mai importante aspecte privind crearea și consultarea unei baze de date terminologice.

Dacă utilizatorul dorește să definească el însuși structura bazei terminologice, are la dispoziție mai multe posibilități:

- a) Își poate defini propria structură bazei de date, după dorință, beneficiind de ajutorul unui asistent [*wizard*] care îi dă posibilitatea să utilizeze tipare predefinite pentru baze terminologice bi- sau multi-lingve, pe care apoi le poate modifica și adapta cerințelor proprii. Utilizatorul are și posibilitatea de a-și defini propria structură a bazei terminologice, prin definirea categoriilor de date dorite.
- b) Poate să utilizeze structura a unei baze terminologice existente, creând o bază nouă, în care să-și încarce proprii termeni.
- c) Poate, de asemenea, să plece de la un glosar terminologic creat în **Excel** și să obțină o structură de bază terminologică similară, în care să importe intrările deja

existente în glosar. Pentru aceasta, însă, trebuie făcută o conversie prealabilă a foi de calcul Excel cu utilitarul **MultiTerm Convert**.

După ce structura bazei de date a fost definită, se pot introduce intrări terminologice. Datele pot fi și importate din fișiere externe (alte baze de date, foi de calcul Excel).

Posibilitățile de interogare a bazei terminologice sunt multiple și flexibile. Căutarea se poate face după orice limbă considerată ca sursă. Dacă sunt mai multe baze terminologice deschise, căutarea se poate face simultan în toate acestea.

Sunt accesibile diverse criterii de căutare: potrivirea exactă, căutarea „fuzzy” (când se localizează toți termenii care conțin șirul de căutat), aplicarea filtrelor.

Direcția de căutare poate fi inversată: dacă s-a plecat de la o limbă sursă A spre o limbă țintă B, prin simpla apăsare a unui buton se poate face căutarea de la limba B spre limba A.

Introducerea unei intrări noi sau editarea uneia existente se fac cu ușurință, în fereastra de introducere fiind afișate câmpurile din structura bazei de date.

Concluzii

Pachetul de programe **Trados**, prin facilitățile pe care le oferă, poate fi considerat un model pentru categoria de instrumente de asistare a traducătorilor. El poate fi folosit cu maximă eficiență de către echipele de traducători din cadrul întreprinderilor de traducere; poate fi folosit, de asemenea, de către traducătorii independenți. Nu în ultimul rând, în procedul didactic poate servi ca exemplu de instrumente de asistare a traducerii.

Referințe bibliografice

*** *Trados Tm 6.5 Training Guide*

TRADOS www.Trados.com

Cherata, Sanda 2004 *Aplicații informatice de asistare a traducătorilor*, în **Uniterm**, Universitatea de Vest, Timișoara, Nr. 2/2004

Dennett, Gerald 1995 *Translation Memory: Concept, products, impact and prospects*, Project Report, MSc ISE Part-Time Year 2, South Bank University, School of Electrical, Electronic and Information Engineering, London, 1995