

## Special Issue: Specialised Translation II

# Tecnologia FLOSS per la traduzione

## Disponibilità, applicazione e problematiche

By Peter Sandrini (University of Innsbruck)

### Abstract & Keywords

**English:** The global changes in technology, society and economy have made the use of digital tools a necessity, this holds true even more for the globalized business of translation and localization. Today, a wide variety of free and libre open source software (FLOSS) tools exist for translation, so that is possible to set up a free and independent workstation for the translator allowing the use not just of translation memories, terminology management or alignment tools, concordancers and other specific tools, but also of free standards. FLOSS tools are especially suited for open and collaborative environments, thus also for research and training at universities. This paper gives an overview over the state of the art in FLOSS translation tools, describes the FLOSS proposals made available by the University of Innsbruck and argues for the adoption of FLOSS translation tools in the academia.

**Italian:** I cambiamenti globali che hanno interessato la tecnologia, la società e l'economia hanno fatto dell'utilizzo degli strumenti digitali una necessità, e questo è tanto più vero per l'attività globalizzata della traduzione e della localizzazione. Oggigiorno, è disponibile un'ampia varietà di applicazioni software liberamente disponibili (FLOSS) per la traduzione. Ciò rende possibile costruire una stazione di lavoro gratuita e indipendente per il traduttore il quale potrà in tal modo disporre non solo di memorie di traduzione, di strumenti di gestione della terminologia o di allineamento, di concordancers e di altri strumenti specifici, ma anche di standard gratuiti. Gli strumenti FLOSS si prestano particolarmente ad ambienti aperti e collaborativi e pertanto anche alla ricerca e alla formazione presso le Università. Il presente lavoro offre una panoramica dello stato dell'arte degli strumenti FLOSS per la traduzione, ne descrive le proposte elaborate dall'Università di Innsbruck e propone l'adozione di tali strumenti in ambito accademico.

**Keywords:** computer-aided translation, translation technology, free software, open source, traduzione assistita, tecnologia della traduzione, software libero

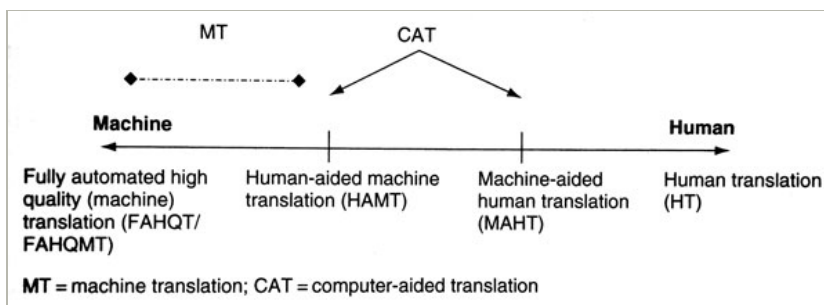
### 1. Introduzione

Molto è cambiato nella professione del traduttore durante gli ultimi due decenni: a partire dagli sviluppi tecnologici nell'informatica fino ad arrivare all'ubiquità della grande rete. Meno di trenta anni fa i traduttori lavoravano esclusivamente con carta e penna oppure con la macchina da scrivere senza alcuna necessità di usare applicazioni software. Successivamente, il personal computer ha cambiato radicalmente il quadro: a partire dagli ultimi anni ottanta del secolo scorso, tutti lo usavano con le applicazioni software installate da dischetti o cd-rom. Meno di un altro decennio dopo, verso la metà degli anni novanta, la seconda grande rivoluzione ebbe inizio: la grande rete con il WWW fece sì che nessuno oggi possa fare a meno di un dispositivo digitale con cui accedere alla rete ed ai suoi innumerevoli servizi. Dispositivi che possono essere un cellulare di ultima generazione oppure un computer portatile o fisso – tutti comunque hanno bisogno di software specifico, non solo per accedere alla rete, ma anche per poter svolgere il proprio lavoro.

Ciò vale anche per la professione del traduttore: oggi, molti tipi di traduzioni non sarebbero possibili, oppure possibili soltanto in modo limitato senza un adeguato supporto informatico, si pensi alla localizzazione del software oppure alla traduzione dei siti web. Non solo, anche l'enorme crescita del mercato della traduzione a seguito della globalizzazione indotta dalla rete ha fatto sì che le traduzioni dovevano essere svolte in sempre meno tempo ed in un numero sempre maggiore di lingue; vedasi ad esempio il concetto di "simship" (simultaneous shipping) dell'industria del software: il lancio simultaneo e parallelo di tutte le versioni linguistiche di un determinato prodotto. Questo comporta ovviamente la necessità di poter aumentare la produttività della traduzione, un lavoro che come tutti noi sappiamo richiede una lunga ed onerosa preparazione individuale. Quindi, se si vuole aumentare l'output dei traduttori a parità di livello qualitativo, non resta altro che sfruttare tutte le possibilità offerte dalla tecnologia digitale. Vediamo dunque, quale può essere l'apporto della tecnologia moderna della traduzione.

### 2. Tecnologia della traduzione

Molti termini sono stati creati per descrivere l'impatto dell'era digitale sulla traduzione: "traduzione assistita dal computer" ovvero "computer aided translation CAT", "traduzione automatica assistita" (in inglese "machine aided human translation MAHT") oppure il termine più recente "human translation technology HTT" usato in ambito UE. Hutchins e Somers (1992: 148) hanno proposto una prima classificazione delle attività di traduzione in base al supporto tecnologico fornito:



Classificazione delle attività di traduzione elaborata da Hutchins e Somers (1992)

Con l'avvento di Internet e la possibilità di combinare tutti i tipi di traduzione assistita nonché i vari scenari di condivisione di risorse linguistiche in rete e le forme di collaborazione tra gruppi di traduttori richiesta e perfino resa necessaria dai grandi progetti di localizzazione, questa suddivisione netta non era più sostenibile.

Noi preferiamo, quindi, il termine "tecnologia della traduzione" usato ad esempio da Quah (2006), un concetto più moderno ed ampio della traduzione assistita in quanto comprende anche tutti gli strumenti al di fuori del processo traduttivo vero e proprio, come ad esempio la traduzione automatica, gli strumenti informatici per la localizzazione del software e la gestione della terminologia nonché per la ricerca delle concordanze in corpora testuali ecc., senza imporre una distinzione netta tra i vari tipi.

Melby (2000) parlando proprio della translation technology distingue diversi tipi di applicazioni:

La base è costituita dall'infrastruttura digitale che comprende tutto il necessario per poter gestire una postazione di lavoro del traduttore: sistema operativo, supporto delle lingue e delle scritture diverse, elaborazione testi oppure suite per ufficio con supporto linguistico, accesso ad Internet con browser e posta elettronica, gestione file ed archivi.

Seguono le varie applicazioni per

- la preparazione della traduzione a livello lessicale: la ricerca della terminologia, l'estrazione dei termini, ecc.
- la preparazione della traduzione a livello di segmento: la segmentazione del testo, l'allineamento del testo di partenza e del testo di arrivo, ecc.
- il processo di traduzione a livello lessicale: ricerca ed inserimento automatico della terminologia, creazione automatica di schede terminologiche, ecc.
- il processo di traduzione a livello di segmento: ricerca nelle memorie di traduzione ed inserimento automatico delle corrispondenze, creazione automatica della memoria di traduzione, traduzione automatica, ecc.
- la verifica della traduzione a livello lessicale: verifica della consistenza terminologica, controllo termini obsoleti, ecc.
- la verifica della traduzione a livello di segmento: controllo segmenti mancanti, controllo formattazione, controllo ortografico, ecc.
- la gestione del work-flow, dei progetti e della contabilità nonché della fatturazione.

Un traduttore esperto non solo deve ora conoscere tutti gli strumenti informatici utili per il suo lavoro, ma deve anche essere informato su tutto quello che la tecnologia digitale può offrire nel campo della traduzione. La figura professionale del futuro non sarà più il traduttore solitario bensì l'esperto per la traduzione assistita da opportuni strumenti anche automatici.

La tecnologia della traduzione dunque fa parte del bagaglio formativo del traduttore professionale, e quasi tutti i corsi formativi universitari ne prendono atto. Non sempre purtroppo ciò si riflette sui contenuti ed i percorsi della ricerca, in quanto da un lato gli sviluppatori delle applicazioni di traduzione assistita nonché di traduzione automatica non si sono occupati molto dei risultati della ricerca traduttologica ("there was no evidence of translation theory being used in the development of machine translation, human-aided machine translation or computer-aided translation"; Quah 2006: 176); dall'altro lato anche la traduttologia ha dato poco spazio finora all'impatto delle tecnologie moderne: come ad esempio i problemi derivanti dalla ripartizione del testo in singoli segmenti a livello di frase o di proposizione con le ripercussioni sull'interpretazione del testo di partenza (Yamada 2011, García 2006), la produzione del testo di arrivo e le caratteristiche testuali; oppure l'impatto della traduzione automatica gratuita in rete sulla professione del traduttore in generale, sui tipi di testo da tradurre, sui vari livelli di qualità delle traduzioni ed in particolar modo sulla didattica (Pym/Biau Gil 2006). Anche se esistono vari volumi sulle tecnologie moderne per la traduzione (vedasi ad es. Austerlühl 2001, O'Hagan/Ashworth 2002, Bowker 2002, Koehn 2010), la ricerca sul processo traduttivo, sulle decisioni del traduttore e soprattutto la ricerca nell'ambito della traduzione specializzata, dove l'importanza del supporto tecnologico risulta maggiore, hanno toccato l'argomento solo in modo marginale. In Intonti/Todisco/Gatto (2007), Popović (2006) e Rocca Longo/Leproni (2007) non si trova alcun riferimento alle tecnologie informatiche per la traduzione. La produttività del traduttore in seguito all'uso delle tecnologie moderne nonché il ruolo economico della traduzione specializzata manca anche nei volumi pubblicati sulla traduzione specializzata di Scarpa (2001), Stolze (2007) e Rega/Magris (2004). Ed Osimo (2001) nel suo libro sulla traduzione e le nuove tecnologie addirittura non fa cenno alla memoria di traduzione ed ai relativi risvolti sul lavoro del traduttore. Molte sono invece le iniziative di cooperazione tra atenei ed associazioni professionali proprio nel campo della tecnologia: un segno questo che della tecnologia della traduzione si avvertono più i risvolti pratici che non le potenzialità di ricerca.

Con il termine "tecnologia della traduzione" intendiamo quindi una serie di metodologie, applicazioni software e formati standardizzati attorno alla traduzione: le caratteristiche del testo digitale, i formati di testo, il markup e la segmentazione dei testi, l'archiviazione ed il recupero di dati linguistici, l'analisi automatica di testi e corpora linguistici, lo scambio di dati, le memorie di traduzione, il management di dati terminologici, la tipologia delle applicazioni di traduzione assistita, i tipi ed i modelli di traduzione automatica, i cenni di storia sulla traduzione assistita ed automatica.

I formati standard rendono possibile lo scambio dei dati nonché la libera scelta del software; essi sono regolati da norme internazionali (ISO) oppure da proposte pratiche elaborate da gruppi di interesse, come ad esempio lo "Special Interest Group" OSCAR all'interno della LISA (Localization Industry Standards Association) autore dei formati TMX e TBX.

Le applicazioni software possono essere sviluppate su base commerciale da parte di aziende e protette da diritti d'autore o semplicemente dalla discrezionalità dei programmatori, nonché da sistemi di licenze molto restrittive; in questo caso si parla di

sistemi proprietari. Oppure le applicazioni ed i formati standard possono essere il frutto di una cooperazione globale di individui senza scopo di lucro, dove ognuno – sia programmatore che utente – è libero di usare l'applicazione come desidera: il mondo del software libero.

### 3. FLOSS e software libero

Parlando di software libero si possono distinguere diversi approcci riguardo al modello filosofico ed al tipo di licenza usato. Il modello purista rappresentato dalla Free Software Foundation e dalla licenza GNU/GPL (GNU General Public License) si basa su precisi principi filosofici di libertà: il software costituisce un bene pubblico, la cui produzione avviene mediante uno sforzo congiunto e collettivo; questa cooperazione è più importante del diritto d'autore individuale, perché protegge il diritto fondamentale ed etico all'informazione ed alla condivisione e diffusione del sapere.

Le quattro libertà del software libero sono:

1. libertà di eseguire il programma per qualsiasi scopo;
2. libertà di studiare come funziona il programma e adattarlo alle proprie necessità;
3. libertà di ridistribuire copie in modo da aiutare il prossimo;
4. libertà di migliorare il programma e distribuirne pubblicamente i miglioramenti apportati dall'utente, in modo tale che tutta la comunità ne tragga beneficio.

L'accessibilità al codice sorgente è una condizione necessaria per la seconda e quarta libertà.

Il termine "open source" (OSS Open Source Software), coniato da un gruppo di sostenitori del software libero alla fine del secolo scorso, si basa proprio sull'accessibilità del codice dando maggior rilievo ai vantaggi pratici della diffusione del codice sorgente e dello sviluppo cooperativo del software su Internet e sorvolando sugli aspetti filosofici ed etici della Free Software Foundation.

Oggi, frequentemente si parla di "FLOSS" (Free and Libre Open Source Software) creando in questo modo un iperonimo per tutte le applicazioni software liberamente disponibili. L'uso di questo tipo di software offre indubbi vantaggi in confronto al software proprietario. Prima di tutto si guadagna in indipendenza tecnologica ovvero si evita il cosiddetto "lock-in" cioè la dipendenza da un solo produttore di software per l'infrastruttura digitale, per le applicazioni, per gli aggiornamenti, nonché per il supporto. In secondo luogo si ha la possibilità di adattare l'applicazione software ai propri bisogni, all'infrastruttura digitale della propria azienda o organizzazione, estendendo anche la funzionalità dell'applicazione. Inoltre non esiste nessun vincolo al numero di installazioni del software, che per di più può essere ridistribuito a chiunque.

Questa enorme flessibilità d'uso e il costo iniziale minimo portano ad un abbassamento delle barriere di accesso alla tecnologia riducendo in questo modo anche il digital divide, termine inglese che indica il divario esistente tra chi ha accesso alle tecnologie dell'informazione e chi ne è escluso. Come emerge da un sondaggio svolto dall'associazione nazionale tedesca per traduttori ed interpreti (BDÜ – Bundesverband der Dolmetscher und Übersetzer) nel 2001 (vedasi Höcker/Freigang 2003), uno dei motivi perché i traduttori non fanno uso delle memorie di traduzione, è proprio il costo delle applicazioni proprietarie. Il prezzo delle licenze e degli aggiornamenti continui costituisce anche per gli atenei un problema sempre più pressante in tempi di crisi; della situazione in cui si trovano le università riguardo alle applicazioni software necessarie parleremo più avanti.

Un ulteriore fattore a favore del software FLOSS è costituito dall'applicazione coerente dei formati standardizzati e dei formati aperti, un vantaggio molto importante anche nel campo della traduzione.

La causa dei vantaggi elencati nonché di alcuni svantaggi inerenti al software FLOSS si possono individuare nell'approccio fondamentalmente diverso tra il FLOSS ed il software proprietario. Mentre nel mondo del software commerciale assistiamo ad una concentrazione non solo dei produttori ma anche delle applicazioni fino ad arrivare a programmi molto complessi con una funzionalità assai differenziata, nel mondo del software libero si possono trovare una pluralità di singoli progetti con delle applicazioni che in molti casi vengono sviluppate per un determinato scopo con una funzionalità ristretta oppure circoscritta ad una determinata funzione. Inoltre, lo sviluppo cooperativo e molte volte anche globale del software libero rende necessario l'impiego di opportuni strumenti di cooperazione, primo fra tutti l'adozione di formati standard per agevolare lo scambio di informazioni e del codice sorgente. Nella maggior parte dei progetti questo porta anche ad una predisposizione delle applicazioni al plurilinguismo ovvero ad una più semplice localizzazione del software (vedasi il sistema *GETTEXT* con i file PO e POT) ed un conseguente incorporamento rapido di una pluralità di lingue.

Il software proprietario viene regolarmente aggiornato anche se l'utente non è mai sicuro del fatto che l'azienda continui ad esistere nel tempo oppure che non venga acquistata da un altro produttore di software che poi interrompa lo sviluppo del prodotto. Gli aggiornamenti del software libero invece in molti casi sono irregolari e dipendono in gran parte dai programmatori volontari, anche se alcuni progetti con l'aiuto di grandi gruppi di sostenitori sono capaci di offrire aggiornamenti puntuali e regolari, come ad esempio lo sviluppo del *kernel Linux* oppure il progetto *LibreOffice*. Questo supporto costituisce un grande punto di forza del software libero ed open source, sicché per la maggior parte dei progetti esistono nella rete gruppi di discussione, blog ed altri strumenti per aiutare l'utente nell'uso del software. Il supporto commerciale per entrambi i tipi di software, invece, in molti casi è costoso oppure carente.

Vediamo in seguito quali strumenti e potenzialità di sviluppo può offrire il mondo del software FLOSS per la traduzione.

### 4. FLOSS per traduttori

Oggi, il mondo del software FLOSS è in grado di offrire al traduttore tutta l'infrastruttura digitale occorrente. Incominciando dal sistema operativo, il traduttore può avvalersi del sistema operativo libero *Linux* di cui esiste anche una versione appositamente studiata per i traduttori, il *tuxtrans*: ne parleremo più avanti. Il supporto delle diverse lingue e scritture è garantito dall'implementazione della norma Unicode (ISO 10646), e con il programma *LibreOffice* che a sua volta lavora con il formato standardizzato OpenDocument (ISO 26300) esiste una suite per ufficio a pari livello dei concorrenti proprietari. L'applicazione stessa è stata localizzata in più di 90 lingue e per le lingue più grandi esiste anche il supporto linguistico nella forma di dizionari, correttori ortografici e sillabazioni. Per l'accesso ad Internet ci sono il browser *Mozilla Firefox*, per la posta elettronica l'applicazione *Mozilla Thunderbird*, giusto per elencare i programmi più comuni tra le tante proposte del software libero. Per la gestione dei file e degli archivi l'utente può avvalersi dei programmi del sistema operativo *Linux* oppure scegliere tra le varie applicazioni libere esistenti.

Con questa infrastruttura digitale abbiamo una solida base sulla quale costruire la nostra postazione di lavoro del traduttore basata sul software libero.

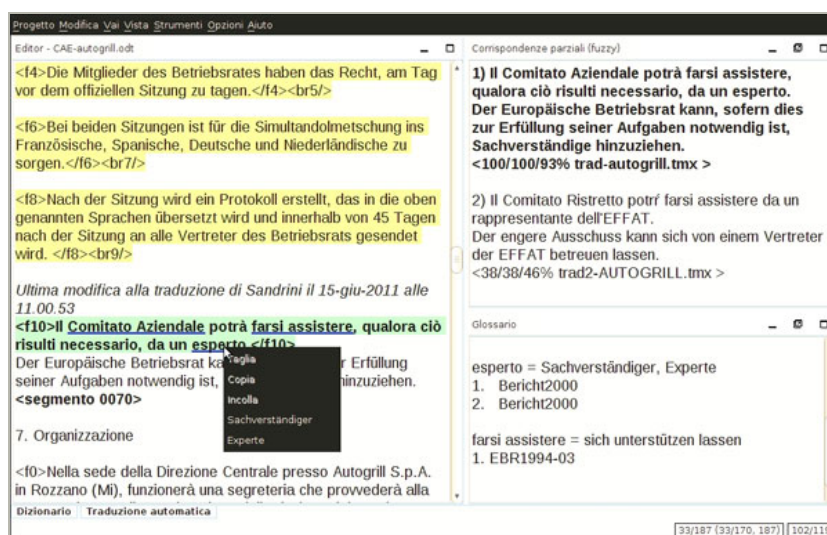
La tendenza alla concentrazione del mercato su pochi produttori che poi raggiungono uno stato di quasi monopolio, la possiamo osservare anche nell'ambito delle applicazioni per la traduzione assistita. L'azienda che nella seconda metà degli anni ottanta del secolo scorso lanciò il primo software per la memoria di traduzione e perciò vanta un'esperienza di più di vent'anni in questo settore, oggi, con un altro nome, è il leader indiscusso con una quota di mercato altissima (vedasi Lagoudaki 2006), con i concorrenti più agguerriti a meno della metà. Molte volte, quindi, il traduttore si vede quasi costretto ad usare un determinato software. Se invece si vuole approfittare dei vantaggi del software libero di cui abbiamo parlato sopra, esistono delle alternative valide. In seguito cerchiamo di fare un confronto esemplare, ma certamente non scientifico e rigoroso, tra un'applicazione commerciale ed uno dei concorrenti FLOSS.

#### 4.1 Le memorie di traduzione

Vediamo con l'esempio delle memorie di traduzione come si presenta il rappresentante FLOSS riguardo al prodotto commerciale più venduto. Per fare questo, dobbiamo per prima cosa analizzare il motivo per il quale questo tipo di applicazioni ha avuto tanto successo nel mondo della traduzione professionale, ovvero la funzionalità di base dei programmi delle memorie di traduzione. Essi sono stati concepiti con l'idea di memorizzare tutte le traduzioni eseguite frase per frase, oppure in modo più tecnico, segmento per segmento, per poter poi controllare nel database delle traduzioni, se all'interno del nuovo testo di partenza da tradurre ci sia un segmento che sia identico oppure somigliante ad un segmento già memorizzato. Il risultato del controllo viene presentato al traduttore che poi decide se usarlo ed inserirlo nella nuova traduzione oppure cambiarlo ed adattarlo alla nuova situazione. In aggiunta, queste applicazioni identificano tutti i termini tecnici presenti all'interno del testo da tradurre se compresi nel database terminologico proponendone l'equivalente nella lingua d'arrivo.

Oltre a queste funzioni basilari, i programmi delle memorie di traduzione offrono anche una ricerca libera nel database delle traduzioni equivalente ad una ricerca delle concordanze, nonché un calcolo sommario delle corrispondenze tra un testo da tradurre ed il contenuto del database arrivando così ad una stima del lavoro da svolgere ed una base di calcolo per un preventivo dei costi. Questa è in sintesi la funzionalità di base delle applicazioni delle memorie di traduzione, attorno alla quale i vari produttori di software hanno costruito i loro prodotti aggiungendo più o meno funzioni secondarie.

Prendendo dal mondo del software libero il programma *OmegaT* e dal mondo commerciale il programma *SDL/Trados Studio 2009*, due applicazioni software molto diverse tra di loro per quanto riguarda la filosofia, la metodologia e la concreta realizzazione, possiamo constatare per prima cosa che in tutti e due i programmi è presente questa funzionalità di base.



#### *OmegaT 2.3.3*

Controllando poi il supporto dei vari formati per i testi digitali da tradurre, possiamo osservare che le differenze non sono tanto grandi: entrambi supportano i file nel formato standardizzato OpenDocument, *Microsoft Office 2007*, XHTML/HTML, i file di solo testo, nonché i file DocBook, un formato XML speciale usato per la documentazione. Entrando però nei dettagli, anche per questi tipi di file molto comuni bisogna differenziare: *OmegaT* ad esempio riesce a leggere ed a salvare i file *Microsoft Office 2007*, incontra però delle difficoltà nella rappresentazione dei codici "in-line" di formattazione. D'altro canto, il filtro di importazione del formato OpenDocument nel programma *SDL/Trados* è molto recente ed ha ancora qualche difficoltà con le formattazioni più complesse. Tutti e due i programmi leggono e scrivono molti altri formati speciali tipici per l'ambiente da cui provengono: *OmegaT* supporta i formati di localizzazione del mondo open source come i file PO, i file MediaWiki, i file JavaProperties, mentre *SDL/Trados* riconosce e scrive molti file provenienti da applicazioni commerciali come i file *Adobe InDesign* e *Framemaker*, i file RTF ed i file DITA.

Per quanto riguarda invece i formati standardizzati, un buon programma per la traduzione assistita dovrebbe attualmente supportare i seguenti formati:

XLIFF (XML Localization Interchange File Format), formato dei file per la localizzazione software capace di contenere anche i testi in più lingue ed altre informazioni utili per i singoli segmenti di testo. Per il software libero viene usato anche il formato PO (Portable Object) e POT (Portable Object Template) meno sofisticato che però può contare su un'infrastruttura solida per la localizzazione dei programmi (il GNU Gettext framework).

TMX (Translation Memory eXchange), formato XML per lo scambio delle memorie di traduzione molto importante per i traduttori che vogliono consegnare le loro memorie di traduzione in un formato neutrale indipendente da un determinato programma.

SRX (Segmentation Rules eXchange), formato per lo scambio delle regole di segmentazione dei testi, importante per aumentare la capacità di riutilizzo delle memorie di traduzione scambiate tra applicazioni diverse.

Le due applicazioni non supportano il formato XLIFF se non in forma ridotta e parziale: *OmegaT* riesce a leggere il formato XLIFF monolingue potendo utilizzare le informazioni linguistiche soltanto attraverso l'uso di *Rainbow*, un altro programma libero studiato appositamente per la conversione dei diversi formati. *SDL/Trados* a sua volta usa un formato XLIFF proprietario e adattato SDLXLIFF leggibile anche da *OmegaT*.

*OmegaT* ed *SDL/Trados* supportano entrambi il formato TMX: *OmegaT* lo usa come formato standard per tutte le sue memorie di traduzione, mentre *SDL/Trados* è capace di esportarlo. Un inconveniente minore è causato dalla diversa codifica dei file TMX (UTF8 e UTF16) nonché dal trattamento diverso degli "in-line codes" ovvero dei codici di formattazione all'interno dei segmenti di testo.

SRX, invece, non è supportato da nessuno dei due programmi, anche se *OmegaT* usa una versione XML per la definizione delle regole di segmentazione facilmente riutilizzabile.

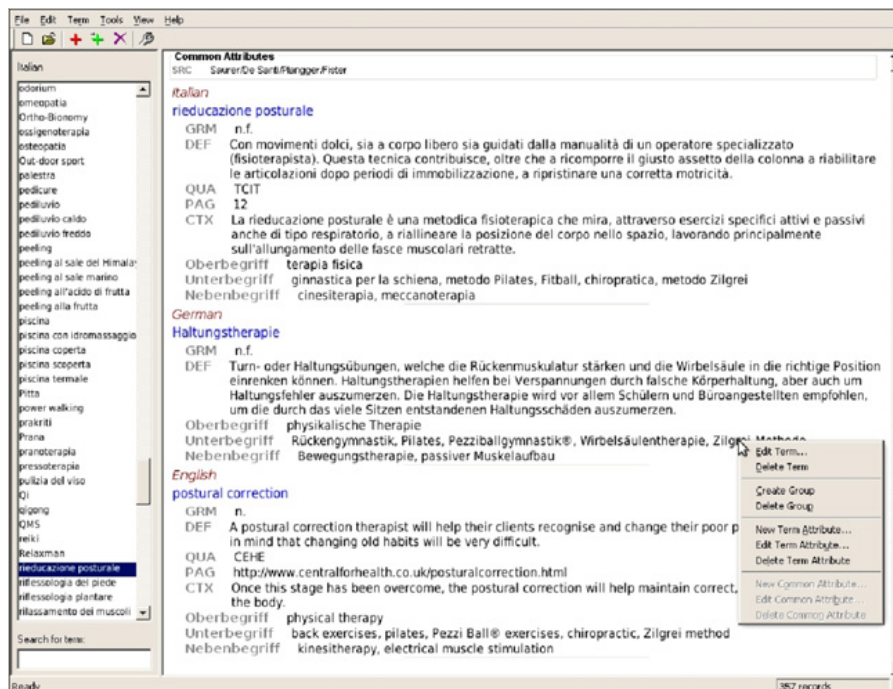
Non volendo entrare nei dettagli delle difficoltà sempre presenti in ogni tipo di scambio dei dati, possiamo constatare comunque che le due applicazioni di memoria di traduzione si equivalgono più o meno nella funzionalità di base e nel supporto dei formati standard. Diversa la situazione se si considerano tutte le funzioni aggiuntive del programma proprietario *SDL/Trados*, quali l'allineamento dei testi, la gestione terminologica, la gestione dei progetti ed il controllo di qualità largamente superiori al concorrente libero.

L'allineamento dei testi costituisce un modulo specifico del programma *SDL/Trados*, *WinAlign* con interfaccia grafica, mentre *OmegaT* non dispone di un modulo di allineamento proprio. Esistono però nel mondo del software libero alcune applicazioni per l'allineamento dei testi come *BiText2TMX*. La differenza più importante è il supporto dei formati di file: *BiText2TMX* legge soltanto i file di solo testo, mentre *WinAlign* legge tutti i formati compatibili con *SDL/Trados*. Per avere la stessa funzionalità bisogna avvalersi di un altro programma di allineamento gratuito, *AlignAssist* software non libero e neanche open source, però messo a disposizione gratuitamente che legge i formati *Microsoft Office*, RTF ed HTML.

#### 4.2 La gestione della terminologia

Per quanto riguarda la gestione della terminologia *SDL/Trados* ha in campo con *Multiterm* un prodotto molto sofisticato e leader incontrastato del mercato già da anni, per cui è stato difficile raggiungere lo stesso livello di qualità e di funzionalità per un progetto open source. Tra le funzionalità di base di un sistema per la gestione della terminologia ci sono da elencare una scheda terminologica completa oppure adattabile dall'utente con i vari campi di informazioni disponibili raggruppati secondo i tre livelli descritti dal Terminology Markup Framework (TMF, norma ISO 16642 e 12620): concetto, lingua, termine. Molto importanti sono anche le funzionalità per lo scambio dei dati con la possibilità di importare ed esportare i dati nei formati più comuni e nel formato standardizzato TBX (Terminology Base eXchange Format). Completano un sistema moderno per la gestione della terminologia, una funzione di ricerca avanzata nonché la gestione di database multipli e la possibilità di gestire la propria terminologia online.

In questo caso una comparazione è presto fatta: quasi tutti questi requisiti non si trovano in un'applicazione libera o open source. Due sono le applicazioni libere a disposizione: la componente terminologica del pacchetto *ForeignDesk Termbase* e *TermWiki*, un programma database per la gestione online della terminologia. Il progetto open source *ForeignDesk Termbase* è fermo nel suo sviluppo più o meno dal 2002, perciò riflette lo stato della gestione terminologica degli anni novanta del secolo scorso: *Termbase* non possiede alcuna gerarchizzazione dei campi di informazioni, solo un raggruppamento dei campi all'interno della scheda terminologica, le possibilità di import ed export sono limitate e non esiste il supporto per il formato TBX.



*ForeignDesk Termbase*

Nonostante tutto, *Termbase* è un programma abbastanza utile che non dovrebbe mancare in una postazione di lavoro libera del traduttore. *TermWiki* è stato lanciato all'inizio del 2010 ed è ancora presto per fare una valutazione definitiva. Si tratta di un'applicazione basata su MySQL, il programma database più diffuso in ambiente server ed Internet, ed è installabile su server dove la gestione della terminologia è demandata a un semplice browser Internet rendendo possibile così una gestione cooperativa

della terminologia senza limiti geografici. La struttura della scheda terminologica di *TermWiki* risulta fissa con la possibilità di definire alcuni campi secondo le proprie esigenze. L'installazione su server MySQL è un compito piuttosto per esperti informatici e non alla portata di un traduttore medio.

#### 4.3 Le differenze tra software libero ed applicazioni proprietarie e commerciali

Il nostro breve raffronto mette a fuoco le differenze fondamentali tra il software libero e le applicazioni proprietarie e commerciali in questione:

da una parte *SDL/Trados*, un pacchetto di programmi completo di memoria di traduzione, gestione della terminologia molto sofisticata, allineamento dei testi di partenza e di arrivo, gestione dei progetti e modulo per il controllo della qualità sulle traduzioni e sulla terminologia usata nel progetto,

dall'altra parte *OmegaT*, un programma per la memoria di traduzione piuttosto puro che raggiunge le sue piene potenzialità se affiancato da una serie di altri programmi liberi: *LibreOffice* e *Rainbow* per la conversione dei file, *BiText2TMX* per l'allineamento, *FD-Termbase* per la gestione della terminologia.

Il vantaggio del costo zero per le applicazioni libere non può nascondere il fatto che esse richiedano un maggior sforzo per l'utente in quanto egli dovrà disporre delle informazioni sui vari programmi esistenti, installarli uno per uno e saperli usare. Proprio questo costituisce – naturalmente oltre la convinzione che la tecnologia della traduzione libera sia un'alternativa valida – uno dei motivi per cui abbiamo cercato di divulgare queste applicazioni con delle iniziative specifiche.

### 5. Proposte FLOSS pronte per l'uso

Supponiamo che il traduttore tipico usi un'applicazione commerciale per la traduzione assistita ed un sistema operativo commerciale senza saper niente del mondo FLOSS. Per offrirgli la possibilità di poter scegliere un'alternativa FLOSS, all'Università di Innsbruck cerchiamo di offrire dei pacchetti software FLOSS che contengano una selezione di quanto può offrire il mondo del software libero per la traduzione.

#### 5.1 USBTrans

Procedendo per fasi, il primo passo verso il mondo FLOSS sarebbe di offrire al traduttore la possibilità di usare e mettere alla prova le singole applicazioni FLOSS senza dover installare o cambiare niente sul suo computer. A questo scopo proponiamo una chiavetta USB con le applicazioni di tecnologia della traduzione già installate e pronte per essere lanciate dalla chiavetta USB, a cui abbiamo dato il nome USBTrans. La chiavetta contiene il pacchetto *LibreOffice* per l'elaborazione dei testi, *OmegaT* per la memoria di traduzione con i vari tokenizer (analizzatori lessicali), *Rainbow* per la conversione dei file, *BiText2TMX* ed *AlignAssist* per l'allineamento dei testi, *FD-Termbase* per la gestione della terminologia, *Cmap lite* e *Freemind* per i sistemi concettuali ed i mindmaps, *AntConc* e *TextStat* per la ricerca delle concordanze, il lettore dei file PDF *Sumatra*, le applicazioni software per la traduzione dei sottotitoli *Subtitle Workshop* e *Jubler*, il software bibliografico *JabRef* e molti altri ancora. Con questa collezione di programmi crediamo di poter dire che un traduttore professionale possa svolgere il proprio lavoro regolarmente senza incontrare troppe limitazioni. E soprattutto, dopo l'uso dei vari programmi, il traduttore può staccare la chiavetta USB e tornare alle sue impostazioni ed ai suoi programmi abituali.

#### 5.2 tuxtrans: Linux for translators

Un secondo passo verso una postazione di lavoro del traduttore esclusivamente basata su software libero consiste nell'uso di un sistema operativo libero completo di applicativi per traduttori: *Linux tuxtrans*. Ebbene, il sistema operativo *Linux*, nato all'incirca vent'anni fa, ha ultimamente fatto dei grandi passi avanti ed oggi dal punto di vista dell'utente non risulta più difficile da gestire rispetto ai sistemi commerciali più diffusi. Dal canto nostro, abbiamo deciso di usare una delle varianti *Linux* più diffuse e facile da usare, *Ubuntu*, su cui costruire un sistema che includa gran parte della tecnologia della traduzione libera, anche se il sistema *Linux* non ha avuto ancora un gran successo tra i traduttori come risulta da un sondaggio specifico (vedasi García González 2008: 9).

Il pacchetto da noi offerto contiene oltre al sistema base *Linux* anche l'applicazione *LibreOffice* e tutte le applicazioni software necessarie per un posto di lavoro digitale moderno: accesso ad Internet con browser e posta elettronica, applicazioni grafiche e multimediali, ecc. Un menu separato elenca tutte le applicazioni software rilevanti per la traduzione: *OmegaT* con i vari tokenizer, *FD-Termbase*, *C-Map*, *BiText2TMX*, *Rainbow*, editori XLIFF, applicazioni per l'analisi dei testi e delle concordanze, sistemi per la traduzione dei sottotitoli, per la localizzazione del software e molti altri programmi ancora. Le applicazioni software sono in gran parte identiche ai programmi installati su USBTrans; per una panoramica degli applicativi disponibili su *Linux* vedasi il sito *tuxtrans* nel menu *Applications* oppure Diaz Fouces (2008: 57).

Il sistema *tuxtrans* è scaricabile in forma di file ISO, un formato che con le opportune applicazioni (ad es. Burnatonce, Nero) si può incidere direttamente su disco DVD oppure salvare su una chiavetta USB con almeno 2 GB di spazio. Si tratta di un sistema "live" che si può provare senza alcuna installazione: basta inserire il disco DVD oppure la chiavetta USB, riavviare il pc e lavorare con il sistema *tuxtrans*. Al termine della prova, dopo un riavvio del sistema senza il DVD o la chiavetta, si torna alla propria vecchia configurazione. Se si decide di installarlo definitivamente su disco fisso – accanto od in sostituzione di Windows oppure MacOSX – lo si può fare facilmente lanciando uno script dal desktop che da avvio alla procedura di installazione definitiva.

#### 5.3 Prospettive

La chiavetta USBTrans ed il sistema *Linux tuxtrans* mettono a disposizione due soluzioni basate su software libero per la postazione di lavoro del traduttore. L'uso della tecnologia della traduzione non si limita però al traduttore professionale oppure volontario singolo. Come già detto, alcuni tipi di traduzioni, soprattutto la localizzazione e la traduzione dei siti web, richiedono il lavoro in gruppo sia a causa della mole di lavoro sia a causa dell'ubicazione geografica dei singoli esperti coinvolti. La rete ed il WWW 2.0 forniscono l'infrastruttura per le applicazioni decentralizzate e cooperative, basata in gran parte su sistemi software liberi, tra cui il database open source più comune in rete MySQL oppure il server Apache ed i linguaggi di scripting PHP, Perl e Python.

Esistono già una serie di applicazioni libere per la traduzione cooperativa e condivisa tra più utenti in rete, citiamo ad esempio *Pootle*, *Tumatxa* per le memorie di traduzione multi-user in rete, *TermWiki* per la gestione della terminologia, la gestione dei progetti di traduzione *ProjectOpen* oppure *Globalsight* ed altri ancora. In sintesi, si tratterebbe di installare un sistema *Linux*

con server LAMP (*Linux*, server HTTP Apache, database MySQL e linguaggi di scripting PHP e Python) con tutte le applicazioni di tecnologia della traduzione libera operanti in rete preconfigurate, compilarne un'immagine ISO e distribuirla pronta per l'uso. La si potrebbe chiamare *CloudTrans*.

Un simile prodotto potrebbe essere utile per tutti i progetti di traduzione cooperativa, sia commerciali che volontari, nell'ambito di agenzie di traduzione oppure all'interno di organizzazioni non governative (ONG) oppure governative internazionali oppure ancora per minoranze linguistiche in cerca di un'infrastruttura tecnica per organizzare terminologia bi- o multilingue e le traduzioni.

## 6. Università e tecnologia FLOSS

Anche se non ci sono dubbi sul fatto che la tecnologia della traduzione costituisca una parte integrante della formazione del traduttore, estensione, metodi e contenuti di questa componente didattica sono oggetto di discussione ed anche fondamentalmente diversi tra i singoli corsi universitari. I progetti LETRAC (Language Engineering for Translation Curricula), LEIT (Language Engineering Initiative Task force) all'interno della LISA e eCoLoRe (electronic Content Localisation Resources) hanno cercato di proporre contenuti e metodi per la traduzione e la localizzazione.

Tuttavia, la costante evoluzione e trasformazione della tecnologia della traduzione comporta un onere finanziario non trascurabile per la maggior parte delle università (Maia 2008: 58). Molti produttori di software proprietario cercano di affrontare questo problema mettendo a disposizione degli istituti di formazione versioni ridotte nel prezzo o addirittura gratuite. Il vantaggio per le case produttrici è ovvio: usando un determinato programma nei corsi di formazione si crea una sorta di legame tra gli studenti ed il prodotto che sicuramente influirà sulle scelte future dei traduttori.

Il pericolo che ne deriva, consiste nel fatto che frequentemente la formazione si esaurisce nella preparazione all'uso di un determinato programma; un obiettivo analogo ai corsi di preparazione offerti dalle stesse case produttrici oppure da centri specializzati di assistenza. In alcuni casi le università si trovano perfino ad organizzare esami di certificazione pensati e svolti dalle aziende.

A questo riguardo ci sono da fare una serie di considerazioni sul ruolo delle università nonché sugli obiettivi della formazione. Gli atenei non sono nati per riprodurre il sapere di aziende e produttori di software riducendo così la loro funzione a puro luogo di intermediazione tra aziende e studenti o addirittura a veicoli pubblicitari. Le università devono essere bensì luogo di ricerca e di creazione di contenuti e di sapere originali.

Con il codice sorgente aperto ed accessibile per verifiche, miglioramenti oppure cambiamenti il software libero offre la possibilità di una partecipazione attiva nello sviluppo del programma contribuendo all'aumento del sapere comune. Gli studenti possono studiare come funzionano le applicazioni software, discutere sui metodi applicati ed articolare proposte di perfezionamento: hanno accesso al sapere. I metodi applicati, il codice sorgente dei programmi commerciali sono invece tenuti segreti, chiusi da licenze restrittive.

Inoltre, tutti i risultati della ricerca devono essere pubblicati e il pubblico deve essere in grado di verificare i risultati anche con l'aiuto di tutti gli strumenti usati dai ricercatori. Se i ricercatori usano applicazioni proprietarie queste sono vincolate da licenze e quindi non accessibili al pubblico.

Oltre a questa libertà di utilizzo, con il software libero le università non sono sottoposte alle solite restrizioni per quanto riguarda il numero di installazioni, oppure le copie per gli studenti. Come per la ricerca, anche per gli strumenti e per l'infrastruttura digitale l'indipendenza da produttori e aziende costituisce un bene etico. Utilizzare software libero, dunque, è soprattutto una scelta di natura etica, poiché il suo sviluppo si basa sugli stessi principi fondanti della comunità scientifica, senza i quali la ricerca non può progredire: *il libero scambio delle informazioni, la condivisione di idee e risultati e il libero utilizzo del patrimonio comune delle conoscenze per un'ulteriore sviluppo*.

Non trascurabile è anche il vantaggio dei costi; si tratta comunque di finanziamenti pubblici da spendere responsabilmente: prima di ogni spesa per una qualsiasi applicazione software si dovrà verificare quali alternative valide esistano ed a quale costo. Infatti, se tutti i finanziamenti erogati dalle università per le licenze di software proprietario nonché il sapere dei ricercatori nelle singole discipline fossero stati messi a disposizione dei progetti di software libero, chissà quali applicazioni libere esisterebbero oggi.

Al posto della preparazione all'uso di un determinato programma, anche se leader incontrastato del mercato, subentra l'insegnamento di metodi, criteri e principi della tecnologia della traduzione ed esemplificazione con l'aiuto di applicazioni libere. Al posto del sapere quale tasto premere per una determinata funzione nell'applicazione proprietaria, gli studenti dovrebbero capire ad esempio come funziona la memoria di traduzione, in quali scenari d'uso la memoria di traduzione porta ad un aumento della produttività del traduttore, quali tipi di testi non sono adatti, ecc. Nell'ambito di corsi per esperti in tecnologia di traduzione si potranno studiare ad esempio gli algoritmi di comparazione dei segmenti di testo, le regole di segmentazione oppure il conteggio delle statistiche confrontandoli tra le varie applicazioni libere concorrenti oppure applicandone nuovi metodi.

L'uso del software libero, comunque, dovrà essere integrato anche nei progetti didattici di traduzione come descritti in dettaglio ad esempio da Cánovas e Samson (2008: 193).

All'Università di Innsbruck il software libero con la chiavetta USBTrans e tutti gli applicativi installati costituisce parte integrante del corso "Introduzione alla tecnologia della traduzione" obbligatorio (30 ore) per gli studenti del master in traduttologia, un'idea che è stata accolta anche all'interno di altri atenei (vedasi PortableCat).

Infatti, l'esperienza ha dimostrato che l'accesso alle applicazioni libere installate sulla chiavetta USBTrans risulta immediato e facile da parte degli studenti, in quanto abituati all'utilizzo di applicazioni in ambiente Windows, mentre sembra che ci siano maggiori ostacoli per quanto riguarda l'utilizzo di *tuxtrans* ovvero del sistema operativo libero Linux. Per questo motivo abbiamo deciso di utilizzare soprattutto la chiavetta USBTrans nel corso introduttivo sulla tecnologia della traduzione, riservando, tuttavia, due ore di lezione a un'adeguata presentazione del sistema *tuxtrans*.

In aggiunta, nel corso degli ultimi anni, l'uso didattico della tecnologia della traduzione libera si è dimostrato particolarmente adeguato al primo approccio alla traduzione assistita, nonché per conoscere i vari tipi di applicazioni ed impararne l'uso. Uno studio approfondito di funzioni attuali richiede però spesso l'uso di programmi commerciali avanzati. Riteniamo, dunque, che l'uso esclusivo di tecnologia della traduzione libera nella didattica della traduzione assistita sia certamente possibile, bensì

limitato ad un livello introduttivo, mentre per i corsi di approfondimento le possibilità di utilizzo dipendono molto dallo stato di sviluppo dei singoli progetti open source. All'Università di Innsbruck, pertanto, offriamo un corso aggiuntivo di approfondimento (30 ore) accanto al corso introduttivo utilizzando anche sistemi di traduzione assistita commerciali. Lo studio del codice sorgente e dei metodi di programmazione usati resta riservato a studenti e docenti adeguatamente preparati nel campo informatico nonché nei diversi linguaggi di programmazione; una barriera molto impegnativa che si può superare soltanto all'interno di un master specifico oppure di una specializzazione nel campo della tecnologia della traduzione.

## 7. Conclusioni

In sintesi, vorremo sottolineare che il software libero e le applicazioni FLOSS costituiscono valide proposte nel campo delle tecnologie della traduzione, anche se tuttora ci sono delle lacune e delle carenze in alcuni ambiti. Con un'impostazione adeguata dei corsi, l'introduzione e l'uso della tecnologia della traduzione libera porterà alle università dei vantaggi cospicui. Viceversa, anche il mondo del software libero potrà avvantaggiarsi di una cooperazione solida tra il mondo accademico ed il mondo del volontariato digitale dei progetti open source potendo contare su un input di conoscenze specifiche ed una base crescente di utenti.

## Bibliografia

- Austermühl, Frank (2001) *Electronic Tools for Translators*. Manchester: St. Jerome Publ.
- Bowker, Lynne (2002) *Computer-aided Translation Technology. A Practical Introduction*. Ottawa: University of Ottawa Press.
- Cánovas, Marcos / Samson, Richard (2008) "Dos Ejemplos de Aplicación del Software Libre en la Docencia de la Traducción". *Traducir con software libre*, ed. da Oscar Díaz Fouces / Marta García González, 193-210.
- Díaz Fouces, Oscar / García González, Marta eds. (2008) *Traducir con software libre*. Granada: Editorial Comares.
- García, Ignacio (2006) "Translators on translation memories\_ a blessing or a curse?" *Translation Technology and its Teaching (with much mention of localization)* ed. da Anthony Pym, Alexander Perekrestenko e Bram Starink. Tarragona: Intercultural Studies Group. 97-105.
- García González, Marta (2008) "Free Software for Translators: Is the Market ready for a Change?" *Traducir con software libre*, ed. da Oscar Díaz Fouces / Marta García González, 9-31.
- Höcker, Mary / Freigang, Karl-Heinz (2003): "Nutzung von Translation-Memory-Werkzeugen. Ergebnisse der Umfrage im Rahmen des eCoLoRe-Projekts". *Mitteilungen für Dolmetscher und Übersetzer*. Heft 3, 19-21.
- Hutchins, William J. / Somers, Harold L. (1992) *An Introduction to Machine Translation*. London: Acad. Press.
- Intonti, Vittoria / Todisco, Graziella / Gatto, Maristella (2007) *La Traduzione. Lo stato dell'arte - Translation. The State of the Art*. Ravenna: Longo Editore.
- Koehn, Philipp (2010) *Statistical Machine Translation*. Cambridge: Cambridge University Press.
- Lagoudaki, Elina (2006): *TM Survey 2006*. London: Imperial College.
- Maia, Belinda (2008) "CORPÓGRAFO V.4. Tools for educating Translators". *Topics in Language Resources for Translation and Localisation*, ed. da Elia Yuste Rodrigo, 57-70.
- Melby, Alan (2000) *Eight Types of Translation Technology*.  
online paper: <http://www.ttt.org/technology/8types.pdf> (01/2005)
- O'Hagan, Minako / Ashworth, David eds. (2002) *Translation-Mediated Communication in a Digital World: facing the challenges of globalization and localization*. Clevedon: Multilingual Matters.
- Osimo, Bruno (2001) *Traduzione e Nuove Tecnologie: Informatica e Internet per traduttori. Guida pratica con glossario*. Milano: Hoepli.
- Popovič, Anton (2006) *La Scienza della Traduzione. Aspetti metodologici. La comunicazione traduttiva*. Milano: Hoepli.
- Pym, Anthony / Perekrestenko, Alexander / Starink, Bram (2006) *Translation Technology and its Teaching (with much mention of localization)*. Tarragona: Intercultural Studies Group.
- Pym, Anthony / Gil, José Ramón Biau (2006) "Technology and Translation. A pedagogical overview" *Translation Technology and its Teaching (with much mention of localization)* ed. da Anthony Pym, Alexander Perekrestenko e Bram Starink. Tarragona: Intercultural Studies Group. 5-19.
- Pym, Anthony (2011) *Translation research projects 3*. Tarragona: Universitat Rovira i Virgili.
- Quah, Chiew Kin (2006) *Translation and technology*. Palgrave textbooks in translating and interpreting. New York: Palgrave Macmillan.
- Rega, Lorenza / Magris, Marella eds. (2004) *Übersetzen in der Fachkommunikation = Comunicazione specialistica e traduzione*. Tubinga: Narr.
- Rocca Longo, Marinella / Leproni, Raffaella eds. (2007) *La Babele Mediatica: multiculturalità e comunicazione*. Bologna: Ed. Kappa.
- Scarpa, Federica (2001) *La Traduzione Specializzata: lingue speciali e mediazione linguistica*. Milano: Hoepli.
- Stolze, Rade Gundis (2007) *Die Fachübersetzung. Eine Einführung*. Tubinga: Narr.
- Yamada, Masaru (2011) "The effect of translation memory databases on productivity". *Translation research projects 3*, ed. da Anthony Pym, 63-70.



Yuste Rodrigo, Elia ed. (2008) *Topics in Language Resources for Translation and Localisation*. Amsterdam: Benjamins.

#### Sitografia

USBTrans (ambiente Windows): <http://www.petersandrini.net/it-fsftrans.html>

tuxtrans – Sistema Linux per Traduttori: <http://www.uibk.ac.at/tuxtrans/>

Free Software Foundation: <http://www.fsf.org>

Free Software Foundation Europe: <http://www.fsf.org>

Archivio open source: <http://www.sourceforge.net>

PortableCat: <http://www.traduccionmundolibre.com/2009/12/16/portablecat/>

OpenOffice: <http://www.openoffice.org>

OmegaT: <http://www.OmegaT.org>

Rainbow: [okapi.sourceforge.net/](http://okapi.sourceforge.net/)

ForeignDesk: [foreigndesk.sf.net](http://foreigndesk.sf.net)

Distribuzioni Linux: <http://www.distrowatch.com>

Ubuntu: <http://www.ubuntu.com>

LETRAC: <http://www.iai-sb.de/forschung/content/view/37/50/>

eCoLoRe: <http://www.leeds.ac.uk/cts/en/research.htm>

Floss Manuals: [en.flossmanuals.net/open-translation-tools/](http://en.flossmanuals.net/open-translation-tools/)

©inTRAlinea & Peter Sandrini (2011).

"Tecnologia FLOSS per la traduzione", *inTRAlinea* Special Issue: Specialised Translation II.

This article can be freely reproduced under [Creative Commons License](http://creativecommons.org/licenses/by/3.0/).

**Permalink:** <http://www.intraline.org/specials/article/1796>



[Previous](#) [Next](#)

