



**UNIVERSITÀ DEGLI STUDI DI PALERMO**

Dottorato in Letterature moderne e studi filologico - linguistici

Dipartimento di studi umanistici

L/LIN 01

**LA DIMENSIONE COGNITIVA  
NELLA TRADUZIONE ASSISTITA DA COMPUTER  
E NELLA TRADUZIONE AUTOMATICA**



**IL COORDINATORE  
PROF.SSA FLORA DI LEGAMI**

**IL DOTTORE  
MARILENA RONDELLO**

**IL TUTOR  
PROF.SSA M. LUCIA ALIFFI**

CICLO XXIV – 2011-2013  
ANNO CONSEGUIMENTO TITOLO 2015

DOTTORATO





## Ringraziamenti

In maniera semplice, ma doverosa, voglio ringraziare tutti coloro che mi hanno sostenuto, incoraggiato e aiutato durante questi anni. In particolare, un immenso grazie alla mia famiglia, nel senso più ampio ed inclusivo del termine, per l'affetto e la fiducia dimostratami, e per la presenza costante.

Un sentito ringraziamento con stima e riconoscenza va anche alla prof.ssa Lucia Aliffi e al prof. Antonino Disparti, per essere stati delle ottime guide.

*A mio figlio, Gabriele*

## Indice

<u>INTRODUZIONE</u> .....	pag. 10
---------------------------	---------

### PARTE PRIMA – CAT TOLS E SUSSIDI DIGITALI PER LA TRADUZIONE

#### Capitolo 1: Una panoramica generale

<i>1.1 La nascita dei sussidi digitali per la traduzione</i> .....	pag. 19
<b>1.1.1. La rivoluzione tecnologica</b> .....	pag. 19
<b>1.1.2. Gli esordi</b> .....	pag. 21
<b>1.1.3. I primi CAT Tools</b> .....	pag. 24
<i>1.2 L'impatto della tecnologia sul lavoro del traduttore</i> .....	pag. 25
<b>1.2.1. Tecnologia e traduzione</b> .....	pag. 25
<b>1.2.2. Esigenze comunicative e organizzative</b> .....	pag. 26
<b>1.2.3. I tempi di lavoro e le agenzie</b> .....	pag. 27
<i>1.3 L'industria della traduzione</i> .....	pag. 28
<b>1.3.1. Una descrizione</b> .....	pag. 29
<b>1.3.2. Il fenomeno della localizzazione</b> .....	pag. 30
<b>1.3.3. Localizzazione vs. traduzione</b> .....	pag. 31
<b>1.3.4. Siti web e videogames</b> .....	pag. 33
<i>1.4 Studi linguistici sui CAT Tools</i> .....	pag. 35
<b>1.4.1. Gli Applied Translation Studies</b> .....	pag. 35
<b>1.4.2. Limitatezza degli studi</b> .....	pag. 37
<b>1.4.3. La linguistica computazionale</b> .....	pag. 39
<b>1.4.4. La linguistica dei corpora</b> .....	pag. 40
<b>1.4.5. Lo <i>specialized discourse</i></b> .....	pag. 43

#### Capitolo 2: Tipologie di strumenti

<i>2.1 Translation Memories</i> .....	pag. 45
<b>2.1.1. Le TM</b> .....	pag. 45

<b>2.1.2. Fuzzy match, Exact match</b> .....	pag. 46
<b>2.1.3. L'allineamento</b> .....	pag. 49
<b>2.1.4. Sistemi alternativi</b> .....	pag. 51
2.2 <i>L'ambiente di lavoro</i> .....	pag. 52
<b>2.2.1. La Workstation</b> .....	pag. 52
<b>2.2.2. Le funzioni disponibili</b> .....	pag. 55
2.3 <i>Strumenti di gestione terminologica</i> .....	pag. 57
<b>2.3.1. I terminology management tools</b> .....	pag. 58
<b>2.3.1. I terminology extraction tools</b> .....	pag. 61
2.4 <i>Strumenti per la localizzazione e il project management</i> .....	pag. 62
<b>2.4.1. Strumenti per la localizzazione</b> .....	pag. 62
<b>2.4.2. Strumenti per il project management</b> .....	pag. 63
<b>2.4.3. Le suite SDL Trados</b> .....	pag. 64
2.5 <i>Le risorse disponibili su internet</i> .....	pag. 67
<b>2.5.1. Le pagine dedicate alla terminologia</b> .....	pag. 67
<b>2.5.2. Le pagine dedicate ai traduttori</b> .....	pag. 70
Capitolo 3: CAT Tools nel processo traduttivo	
3.1 <i>Dentro il processo traduttivo</i> .....	pag. 74
<b>3.1.1. La traduttologia cognitiva</b> .....	pag. 74
<b>3.1.2. Il modello di Kiraly</b> .....	pag. 75
<b>3.1.3. Ricadute cognitive dei CAT Tools</b> .....	pag. 80
3.2 <i>Definizione di una procedura operativa</i> .....	pag. 81
<b>3.2.1. I CAT Tools strumenti di organizzazione del lavoro</b> .....	pag. 82
<b>3.2.2. Collaborazione uomo - computer</b> .....	pag. 83
<b>3.2.3. Impatto sui tempi di lavoro</b> .....	pag. 85
<b>3.2.4. Le impressioni degli utenti</b> .....	pag. 86
3.3 <i>Scomposizione del testo</i> .....	pag. 87
<b>3.3.1. La segmentazione frase per frase</b> .....	pag. 87
<b>3.3.2. L'unità cognitiva di traduzione</b> .....	pag. 89
<b>3.3.3. Studi sull'argomento</b> .....	pag. 91
3.4 <i>La memoria, le memorie</i> .....	pag. 93
<b>3.4.1. Le diverse memorie</b> .....	pag. 94

<b>3.4.2. Memoria e traduzione</b> .....	pag. 96
<b>3.4.3. La memoria di lavoro a lungo termine</b> .....	pag. 97
<b>3.4.4. L'influenza dei CAT Tools</b> .....	pag. 99
3.5 <i>Pause come indicatori di attività cognitiva</i> .....	pag. 100
<b>3.5.1. Funzione delle pause nel processo traduttivo</b> .....	pag. 101
<b>3.5.2. L'impatto dei CAT Tools</b> .....	pag. 102
3.6 <i>Conclusioni</i> .....	pag. 104
Capitolo 4: CAT Tools strumento e oggetto della didattica	
4.1 <i>Apprendere una competenza professionalizzante</i> .....	pag. 107
<b>4.1.1. La dimensione didattica</b> .....	pag. 107
<b>4.1.2. L'approccio costruttivista e la centralità del laboratorio</b> .....	pag. 109
<b>4.1.3. L'apporto dei CAT Tools nella didattica</b> .....	pag. 112
4.2 <i>A sostegno dell'acquisizione di una competenza traduttiva</i> .....	pag. 114
<b>4.2.1. In cerca di una definizione</b> .....	pag. 114
<b>4.2.2. Conoscenze tecniche e competenza traduttiva</b> .....	pag. 118
<b>4.2.3. Alcuni studi ed indagini</b> .....	pag. 119
4.3 <i>CAT Tools strumenti di analisi linguistica</i> .....	pag. 121
<b>4.3.1. Un confronto con i corpora</b> .....	pag. 121
<b>4.3.2. Funzioni di ricerca</b> .....	pag. 124
4.4 <i>A sostegno dell'apprendimento dei linguaggi specialistici</i> .....	pag. 128
<b>4.4.1. I Languages for Special Purposes</b> .....	pag. 128
<b>4.4.2. Lo studio tramite i corpora e i CAT Tools</b> .....	pag. 129
4.5 <i>L'esperienza dell'Università di Palermo</i> .....	pag. 132
<b>4.5.1. La prevalenza dell'approccio pratico</b> .....	pag. 132
<b>4.5.2. L'esperienza della Summer School</b> .....	pag. 133
<b>4.5.3. I laboratori di CAT Tools</b> .....	pag. 136

## PARTE SECONDA – LA TRADUZIONE AUTOMATICA

### Capitolo 5: La storia e i modelli

5.1 <i>Una introduzione</i> .....	pag. 141
<b>5.1.1. Un argomento controverso</b> .....	pag. 141
<b>5.1.2. Le voci contrarie</b> .....	pag. 143

<b>5.1.3. Prospettive diverse</b> .....	pag. 144
5.2 <i>Sistemi di prima generazione</i> .....	pag. 145
<b>5.2.1. I primi studi</b> .....	pag. 145
<b>5.2.2 I sistemi <i>direct-approach</i></b> .....	pag. 147
<b>5.2.3. I sistemi <i>phrase-based</i> e il linguaggi controllati</b> .....	pag. 148
<b>5.2.4. Il report ALPAC</b> .....	pag. 151
5.3 <i>Sistemi di seconda generazione</i> .....	pag. 152
<b>5.3.1. L’approccio funzionale</b> .....	pag. 152
<b>5.3.2. Il nuovo approccio formalista</b> .....	pag. 153
<b>5.3.3. Gli approcci <i>transfer-based</i> e <i>interlingua-based</i></b> .....	pag. 156
<b>5.3.4. I sistemi <i>corpus-based</i></b> .....	pag. 157
<b>5.3.5. I sistemi <i>example-based</i></b> .....	pag. 159
5.4 <i>Ostacoli e criticità</i> .....	pag. 161
<b>5.4.1. Risultati parziali</b> .....	pag. 161
<b>5.4.2. Individuare le criticità</b> .....	pag. 162
<b>5.4.3. I limiti materiali</b> .....	pag. 164
Capitolo 6: La traduzione automatica oggi	
6.1 <i>Modelli classici e ibridi</i> .....	pag. 166
<b>6.1.1. Un’offerta differenziata</b> .....	pag. 166
<b>6.1.2. Sistemi personalizzati</b> .....	pag. 167
<b>6.1.3. I modelli ibridi</b> .....	pag. 168
<b>6.1.4. Architetture possibili</b> .....	pag. 170
6.2 <i>Traduzione automatica e CAT Tools</i> .....	pag. 172
<b>6.2.1. Un contributo reciproco</b> .....	pag. 173
<b>6.2.2. Alcuni studi sull’argomento</b> .....	pag. 175
6.3 <i>La traduzione automatica sul web – sistemi a confronto</i> .....	pag. 177
<b>6.3.1. Le origini</b> .....	pag. 177
<b>6.3.2. Sistemi a confronto</b> .....	pag. 179
<b>6.3.3. Traduzioni a confronto</b> .....	pag. 183
6.4 <i>Valutazione dei sistemi di traduzione automatica</i> .....	pag. 185
<b>6.4.1. Punti di vista diversi</b> .....	pag. 185
<b>6.4.2. Metodologie di valutazione umana</b> .....	pag. 187

<b>6.4.3. Metodologie di valutazione automatizzata</b> .....	pag. 188
Capitolo 7: Traduzione umana e traduzione automatica	
<i>7.1 Il punto di vista dei traduttori</i> .....	pag. 191
<b>7.1.1. Il questionario d'indagine</b> .....	pag. 191
<b>7.1.2. Il rifiuto della TA</b> .....	pag. 193
<b>7.1.3. La percezione delle tecnologie</b> .....	pag. 194
<b>7.1.4 La TA strumento per i traduttori?</b> .....	pag. 195
<i>7.2 Applicazioni diverse per la TA</i> .....	pag. 196
<b>7.2.1. Obiettivo qualità</b> .....	pag. 196
<b>7.2.2. Utilizzi alternativi</b> .....	pag. 199
<i>7.3 Una prospettiva cognitiva</i> .....	pag. 201
<b>7.3.1. I modelli teorici di riferimento</b> .....	pag. 201
<b>7.3.2. A imitazione dell'uomo</b> .....	pag. 203
<b>7.3.3. Sistemi semi-automatici</b> .....	pag. 204
<b>7.3.4. La percezione degli utenti</b> .....	pag. 206
<i>7.4 Che cosa prospetta il futuro?</i> .....	pag. 207
<b>7.4.1. Ancora molta strada da fare</b> .....	pag. 207
<b>7.4.2. Le lingue minoritarie</b> .....	pag. 209
<b>7.4.3. Altre sfide</b> .....	pag. 210
<i>7.5 Conclusioni</i> .....	pag. 211
<u>CONCLUSIONI GENERALI</u> .....	pag. 215
<u>APPENDICE</u> .....	pag. 219
<u>BIBLIOGRAFIA</u> .....	pag. 220

## Introduzione

La presente tesi di dottorato è dedicata allo studio critico delle diverse tipologie di sussidi digitali per la traduzione e della traduzione automatica. Entrambe queste realtà sono ormai da lungo tempo parte integrante del panorama relativo alla così detta *industria della traduzione*, un settore del mercato che ha vissuto un'espansione notevole grazie alla diffusione dei mezzi di comunicazione di massa e in particolare di Internet, che ha aperto ad aziende, società, enti o singoli individui la possibilità di proporsi ad una clientela mondiale e di raggiungere facilmente interlocutori in tutte le parti del pianeta. Il bisogno di traduzioni è divenuto anche bisogno di specializzazione, con la crescente richiesta di documenti tecnici o dall'elevato grado di settorialità, che ha portato i traduttori a specializzarsi in ambiti ben precisi e a dover fare ricorso delle promettenti tecnologie informatiche per ottenere un supporto nel processo di traduzione, in particolare nel garantire esattezza e costanza terminologica, e relativamente alla possibilità di utilizzare materiale già tradotto. Oltre alle tecnologie di supporto, si andarono diffondendo anche i sistemi automatizzati, che accesero gli entusiasmi degli studiosi nel campo delle intelligenze artificiali e le speranze di coloro ambivano ad ottenere traduzioni in tempi immediati e senza il dispendioso ricorso ai traduttori professionisti. Il panorama del settore della traduzione è stato profondamente trasformato dall'ingresso e affermazione di tali strumenti, e la trasformazione ha avuto esiti sia positivi che negativi, nonché ricadute significative sulle prassi operative dei traduttori, sulle aspettative degli utenti, sulla didattica della traduzione e in alcuni casi sul modo stesso in cui viene percepito il processo che sta dietro alla produzione di una traduzione. Nel corso della tesi verrà concesso spazio tanto alla descrizione degli apporti positivi delle tecnologie, quanto all'identificazione delle loro ricadute negative sul mercato, sui processi, sulla traduzione e sul traduttore stesso.

Esistono numerosi punti di vista dai quali osservare le tecnologie applicate al linguaggio e in particolare alla traduzione. Ad esempio, è possibile osservare questo

fenomeno dalla prospettiva degli sviluppatori delle tecnologie stesse che lavorano prevalentemente all'intero di società private, oppure si può adottare il punto di vista di quei ricercatori legati all'ambito universitario e interessati prevalentemente allo studio delle diverse tipologie di sistemi, spesso tramite progetti congiunti con società esterne. Infine, ed è questo il punto di vista che per forza di cose e per scelta ho fatto mio, si può guardare alle tecnologie per la traduzione dalla prospettiva di quegli studiosi il cui background è di tipo linguistico, e che pensano alla traduzione – a qualsiasi tipo di traduzione – come a un fenomeno di comunicazione umana (Ulrych, 1997: xii) e per giunta uno dei più complessi, che implica l'armonico intreccio di competenze e conoscenze diverse, e il saper cogliere l'utilizzo della lingua all'interno di un contesto socioculturale. Come suggerisce Pym (2011b), tuttavia, la ricerca umanistica non è né può essere indifferente o disconoscente dell'esistenza delle *translation technologies*, dei loro grandi avanzamenti e del modo in cui il loro utilizzo ha pervaso il campo della traduzione tecnico-scientifica. Essa deve cercare di integrare questo nuovo aspetto e il suo apporto all'interno degli studi in traduzione facendo attenzione a mantenere comunque il proprio punto di vista. La ricerca umanistica non può smettere di essere tale, non può diventare studio matematico o ricerca sui sistemi operativi. Essa può trarre da altre discipline spunti, suggerimenti e persino strumenti di analisi, da utilizzare per sviluppare una propria ricerca, che guardi ai fenomeni continuando a mettere al centro l'uomo e i suoi processi mentali, culturali e linguistici.

In particolare, il lavoro affronta l'argomento adottando un punto di vista che in Italia non è stato sinora sviluppato, e che mette in evidenza la dimensione cognitiva legata all'implementazione e all'utilizzo delle risorse digitali per la traduzione. Per "dimensione cognitiva" intendo il modo in cui gli utenti percepiscono questi strumenti, il modo in cui il loro utilizzo trasforma delle prassi operative già consolidate, l'influenza che essi esercitano sui processi traduttivi naturali come la capacità decisionale o la spontanea suddivisione del testo per la lettura e la successiva riproduzione nella lingua di arrivo. Si intende anche il modo e i criteri secondo i quali gli strumenti sono stati sviluppati per cercare di riprodurre quei meccanismi umani ai quali essi intendono venire in soccorso, e di identificare l'idea di linguaggio e di competenza linguistica che ha guidato i loro sviluppatori nel definire le modalità e i criteri operativi. L'idea di perseguire tale tipo di indagine è scaturita dalla mia personale esperienza in qualità di traduttrice nei settori delle

energie rinnovabili e della moda, che mi ha portato a sperimentare in prima persona la richiesta di competenze informatiche specifiche e dell'utilizzo delle tecnologie per la traduzione come una *condicio sine qua non* per l'avviamento di collaborazioni professionali con agenzie di traduzione sia nazionali che internazionali. Ha contribuito anche il mio interesse circa le tecnologie applicate al linguaggio ed in particolar modo alla traduzione. Tale interesse è stato alimentato dalla partecipazione, nel 2008, alla prima edizione della *Summer School "CAT Tools – Localization, internationalization and globalization"* organizzata dall'Università di Palermo, durante la quale ho potuto conseguire la certificazione SDL di primo livello e in seguito quella di secondo livello, attestanti la competenza nell'utilizzo di SDL Trados, uno dei principali e più diffusi software di traduzione assistita attualmente sul mercato. Nei due anni seguenti ho partecipato alle successive edizioni della Summer School in qualità di tutor, tenendo corsi di preparazione alla certificazione di primo livello e potendo osservare direttamente il comportamento degli studenti nei confronti dei CAT Tools (*Computer-Aided Translation Tools*) e delle altre risorse digitali per la traduzione. In seguito, sempre all'interno dell'Università di Palermo ho avuto la possibilità di occuparmi di alcuni corsi laboratoriali incentrati sull'utilizzo dei CAT Tools, potendo sperimentare un'applicazione di più ampio respiro di tali strumenti, non solo come meri sussidi per la traduzione, ma anche come supporti per la formazione linguistica del traduttore. Tale esperienza pluriennale ha consentito un'osservazione ravvicinata del modo in cui gli studenti o i professionisti si accostano all'uso di tali strumenti e di come questi ultimi esercitino un'influenza sui loro atteggiamenti nei confronti del lavoro e sulle loro prassi operative, in maniera positiva o negativa in base alla persona.

All'interesse nei confronti dei CAT Tools si è unito strada facendo anche quello relativo alla traduzione automatica, tema che all'inizio avevo escluso volendo dare la priorità ai meccanismi e ai processi naturali messi in atto dall'uomo nel tradurre. Tuttavia, una trattazione sulle tecnologie applicate alla traduzione che escludesse del tutto la traduzione automatica sarebbe stato mancante di uno dei suoi aspetti più rilevanti e maggiormente discussi in ambito universitario e di ricerca, nonché del riferimento ad uno dei fenomeni più controversi nel panorama dell'industria della traduzione. Volendo mantenere un taglio simile a quello adottato nella discussione sui CAT Tools, ho scelto di approfondire un aspetto particolare della traduzione automatica, ovvero le diverse

concezioni di linguaggio e di competenza traduttiva che soggiacciono alla creazione dei diversi modelli susseguirsi nel corso degli anni, e il tentativo di riproduzione dei processi traduttivi umani tramite il computer, tipico della ricerca sulle intelligenze artificiali alle quali lo sviluppo dei sistemi di traduzione automatica è strettamente correlato.

La tesi, dopo questa introduzione generale di presentazione del lavoro, si sviluppa in due parti, una prima dedicata alla traduzione assistita da computer, la seconda alla traduzione automatica. Le due tipologie di risorse sono state affrontate separatamente per la diversità degli approcci corrispondenti ad esse: di supporto all'attività umana nel primo caso, e (interamente o parzialmente) alternativo ad essa nel secondo. Tale fondamentale diversità implica chiaramente finalità, processi, risultati e conseguenze differenti, come altrettanto diversi sono gli ambiti di applicazione e le percezioni che i traduttori hanno di entrambe le risorse.

## PARTE PRIMA – CAT TOOLS E SUSSIDI DIGITALI PER LA TRADUZIONE

La Prima Parte, come già accennato, sarà interamente dedicata ai CAT Tools. I primi due capitoli avranno un carattere prevalentemente descrittivo, mentre il terzo ed il quarto introdurranno delle prospettive più specifiche ed originali. Nella fattispecie:

Il primo capitolo presenterà i CAT Tools dal punto di vista storico, tracciando un percorso che andrà dalla nascita di tali strumenti e dei primi studi linguistici sul fenomeno fino ai giorni nostri e agli approcci più moderni allo studio e all'utilizzo dei programmi di traduzione assistita (cf. 1.1). Ci si soffermerà a considerare l'impatto che tali sistemi hanno avuto sul lavoro del traduttore, sia in termini pratici – ad esempio relativamente all'organizzazione o alla gestione dei tempi e delle risorse – che in termini più profondi come nell'esigenza di possedere nuove competenze, di collaborare all'interno di un team e condividere le risorse con i colleghi ecc. (cf. 1.2). Verrà anche dedicato dello spazio alla discussione sul concetto di “industria” della traduzione, alla cui affermazione le tecnologie ed i mezzi di comunicazione di massa hanno dato un contributo determinante consentendo l'abbattimento delle distanze negli scambi commerciali come nei contatti e nelle relazioni. In questo spazio verrà presentata la struttura organizzativa delle agenzie di traduzione, le diverse figure professionali che ivi collaborano, e alcuni aspetti specifici come l'industria della localizzazione, che nata all'interno di quella della traduzione si è poi sviluppata in maniera quasi autonoma, ed altre considerazioni utili a comprendere in

seguito l'applicazione dei CAT Tools nelle diverse fasi del lungo percorso della traduzione (cf. 1.3). Infine, verrà tracciata una panoramica delle maggiori teorie e correnti linguistiche dalle quali i CAT Tools hanno attinto ed alle quali hanno al contempo dato un contributo. In particolare, viene fatto riferimento agli *Applied Translation Studies*, alla linguistica dei corpora, alla linguistica computazionale o linguistica informatica, ed agli studi sul *Natural Language Processing* (cf. 1.4).

Il secondo capitolo scende più nello specifico descrivendo le diverse tipologie di strumenti che vengono raccolti sotto il nome di CAT Tools, come le memorie di traduzione (cf. 2.1), gli ambienti di lavoro integrati (cf. 2.2), gli strumenti per la gestione terminologica (cf. 2.3), le suite per il *project management* e i software per la localizzazione (cf. 2.4). Alcuni tra i più celebri software, sia a pagamento che open source, sono messi a confronto per valutarne l'offerta in termini di funzionalità. Delle immagini accompagneranno la descrizione al fine di mostrare più chiaramente l'applicazione dei vari strumenti. Infine, verrà dedicato dello spazio anche alle risorse disponibili in internet, come glossari e TM gratuiti, forum di discussione per traduttori e pagine dedicate alla raccolta di risorse e informazioni utili nell'ambito della traduzione (cf. 2.5). Sebbene tali strumenti non possano essere considerati a tutti gli effetti CAT Tools, sono comunque risorse utilizzate quotidianamente dai traduttori, congiuntamente ai software CAT.

Il terzo capitolo è uno dei capitoli centrali della tesi, e sviluppa una interessante riflessione sulle ricadute dell'adozione dei CAT Tools a livello profondo, sulle dinamiche e le procedure traduttive, innate o acquisite, messe in atto dai traduttori nello svolgimento delle loro attività. Grazie all'apporto dalla linguistica cognitiva e dalle neuroscienze, nel corso degli anni sono stati sviluppati diversi modelli atti a comprendere i processi mentali legati alla traduzione. È stato qui adottato lo schema di Kiraly (1995) per mostrare su quali livelli e componenti del modello neuro-linguistico del processo traduttivo i CAT Tools esercitino un'influenza e come si inseriscano in qualità di risorse ma anche di organizzatori delle prassi operative (cf. 3.1). Facendo riferimento ai principali studi sull'argomento (Dragsted, 2004 e 2006; Christensen, 2010 e 2011; O' Brien, 2006, 2008, 2009 e 2011; Alves e Liparini Campos, 2009), vengono esaminati gli apporti, talvolta benefici talvolta potenzialmente negativi, degli strumenti di traduzione assistita sulla determinazione del tipo di proceduralità adottata (cf. 3.2), sui processi di scomposizione

e analisi testuale (cf. 3.3), di recupero e immagazzinamento di informazioni nella memoria e di messa in atto delle adeguate strategie traduttive (cf. 3.4). Verrà anche valutata l'influenza dei CAT Tools sulla percezione del lavoro da parte dei traduttori, come conseguenza delle ricadute e dei vincoli che l'adozione di tali strumenti comporta, e verrà fatto riferimento in particolare a un aspetto che si ritrova in diversi degli studi sopra citati, ovvero l'analisi delle pause come rivelatrici di attività cognitiva (cf. 3.5).

Il quarto capitolo della tesi è infine dedicato all'aspetto didattico, nello specifico alla valutazione dei possibili benefici arrecati dall'introduzione dello studio pratico dei CAT Tools nei corsi universitari in traduzione. In questo capitolo converge la mia esperienza maturata nell'ambito della didattica dei sussidi digitali per la traduzione. L'approccio adottato non è esclusivamente quello funzionale all'inserimento nel mercato del lavoro, per cui alcuni atenei inseriscono corsi sui CAT Tools per preparare professionalmente gli studenti fornendo loro l'accesso a delle risorse indispensabili. Pur dando a tale aspetto il giusto peso e considerazione (cf. 4.1), col riferimento all'approccio costruttivista mirante a creare un ambiente di apprendimento pratico in cui viene data priorità all'esperienza, alla collaborazione, alla sperimentazione in prima persona e al *learning by doing*, il capitolo presenta una proposta di applicazione alternativa degli strumenti CAT, come alleati nella didattica per lo sviluppo di una competenza traduttiva integrata e completa (cf. 4.2). La proposta, sperimentata nell'ambito dei corsi svolti all'interno delle due Summer School e dei laboratori di sussidi per la traduzione, e sostenuta da alcuni studi (Pym, 2003; Biau Gil e Pym, 2006; Calvi, 2003; Doherty e Moorkens, 2013) è quella di adoperare tali strumenti non solo nella loro maniera convenzionale e funzionale, ma anche nell'apprendimento e nell'analisi linguistica (cf. 4.3), come se fossero dei corpus in miniatura contenenti stralci di linguaggio reale in due lingue allineate. Grazie alle diverse funzioni disponibili per l'estrazione della terminologia e delle collocazioni, i CAT Tools possono agevolare l'apprendimento dei linguaggi specialistici e supportare la riflessione linguistica utile a formare il traduttore (cf. 4.4). Verrà anche compiuto un confronto tra l'approccio al lavoro di traduzione di studenti e professionisti, e si valuterà l'apporto che i CAT Tools possono avere nell'evoluzione delle competenze degli studenti verso procedure e strategie più mature e professionali. Infine, verrà presentata l'esperienza dell'Università di Palermo, con

riferimento alle tre edizioni della Summer School alla quale si è già fatto riferimento e all'attività didattica nell'ambito dei sussidi per la traduzione (cf. 4.5).

## PARTE SECONDA – LA TRADUZIONE AUTOMATICA

La Seconda Parte della tesi si incentrerà sulla traduzione automatica (TA), dedicando ad essa tre capitoli nei quali verrà tracciato un excursus storico con l'avvicendamento dei differenti modelli e approcci, la realtà odierna del fenomeno caratterizzata dal proliferare delle risorse gratuite sul web, e infine una riflessione critica volta non tanto a riprendere l'eterno e infruttuoso dibattito tra favorevoli e contrari alla traduzione automatica, quanto a definire i limiti entro i quali essa può offrire dei vantaggi e i margini di miglioramento che il futuro può concretamente riservare. Il punto di vista col quale verrà compiuta la trattazione è sempre quello cognitivo: i diversi sistemi di TA sono stati sviluppati a partire da idee e concezioni diverse sul processo traduttivo umano, con l'intenzione di riprodurlo tramite meccanismi automatizzati. Tanto nei modelli *rule-based* quanto in quelli *example-based*, il riferimento di partenza è sempre il processo umano e i comportamenti legati alla comprensione ed elaborazione del linguaggio, studiati e analizzati dalla ricerca sulle intelligenze artificiali e sull'elaborazione del linguaggio naturale.

In particolare, il quinto capitolo sarà principalmente volto a presentare il fenomeno dal punto di vista storico, mettendo particolarmente in luce i presupposti, le ideologie e le finalità soggiacenti allo sviluppo dei diversi modelli di traduzione automatica nel tempo (cf. 5.1), e gli ostacoli e le criticità incontrati (cf. 5.4), fondamentalmente legati alla difficoltà (per molti versi irrimediabile) di riprodurre in maniera meccanica i complessi processi linguistici, la cui natura creativa e le infinite possibilità di utilizzo delle regole sulle quali si basa non sono adatte a essere espresse in maniera matematica o attraverso dei comandi univoci e rigidi. Verrà adottata la divisione classica dei modelli in sistemi di prima (cf. 5.2) e seconda generazione (cf. 5.3), le cui metodologie verranno messe a confronto e paragonate, laddove possibile, al processo traduttivo umano che esse tentano di replicare.

Il sesto capitolo presenterà la realtà della traduzione automatica oggi, i progressi sinora compiuti e i modelli di più recente sviluppo come quelli ibridi (cf. 6.1), che mettono insieme i vantaggi delle diverse tipologie di strumenti al fine di migliorare la

qualità della traduzione prodotta grazie all'uso congiunto di parametri grammaticali e sintattici ed esempi di utilizzo reale desunti da corpora e database di testi bilingui. Per ibridazione si intende anche l'utilizzo congiunto di CAT Tools e TA (cf. 6.2), un'applicazione che ha preso piede più all'interno delle agenzie di traduzione che nella prassi quotidiana dei traduttori. Tale fenomeno ha una duplice manifestazione, con l'introduzione della traduzione automatica all'interno dei software CAT, e l'applicazione all'interno della TA degli stessi algoritmi di estrazione e calcolo del grado di match utilizzati dalle memorie di traduzione all'interno dei CAT Tools. Trovano spazio in questo capitolo anche le risorse disponibili gratuitamente sul Web, come Google Translate, Bing Translation Tool ecc., che verranno messe a confronto per valutarne i risultati e comprendere differenze e analogie nei meccanismi di analisi dell'input e produzione dell'output messi in atto da ciascuna di esse (cf. 6.3). Infine, verranno presentati i principali sistemi di valutazione dei sistemi di TA, i cui parametri di analisi sono interessanti per comprendere i principi che determinano cosa venga inteso per "traduzione automatica di qualità" (cf. 6.4).

Il settimo ed ultimo capitolo della tesi presenterà una riflessione critica sul fenomeno della traduzione automatica, a partire dall'analisi dei risultati di un questionario che è stato somministrato a traduttori professionisti con diversi gradi di esperienza nel settore, al fine di ottenere informazioni sul loro rapporto con la traduzione automatica e le risorse digitali per la traduzione in generale (cf. 7.1). Il questionario ha reso possibile raggiungere una visione d'insieme delle attitudini e dei punti di vista dei diretti utilizzatori degli strumenti descritti, spesso in contrasto con le attitudini e i punti di vista di studiosi, ricercatori e sviluppatori di tali sistemi. Si tenterà di definire in maniera concreta l'apporto reale offerto dalla traduzione automatica, analizzando i diversi bisogni degli utenti che ne fanno uso (cf. 7.2) e tentando di tracciare a partire da tali considerazioni una prospettiva cognitiva che consideri la maniera in cui gli utenti si avvicinano a tale strumento, le loro aspettative e bisogni comunicativi, la differenza tra traduttori e utenti generici nel rapporto con questi strumenti (cf. 7.3). Infine si presenteranno le maggiori sfide per il futuro della TA, il cui lento e limitato miglioramento qualitativo non sembra tuttavia scoraggiare né il suo utilizzo da parte di migliaia di utenti ogni giorno, né gli studi e le sperimentazioni sia in ambito accademico che in centri di ricerca privati (cf. 7.4).

*Marilena Rondello*

## PARTE PRIMA

### CAT TOOLS E SUSSIDI DIGITALI PER LA TRADUZIONE

## Capitolo 1: Una panoramica generale

### *1.1 La nascita dei sussidi digitali per la traduzione*

#### **1.1.1 La rivoluzione tecnologica**

Quello del traduttore è uno dei mestieri più antichi del mondo. Sin dal fiorire delle prime civiltà, la capacità di comprendere e comunicare con altri popoli ha rappresentato un fattore strategico per le relazioni economiche e politiche e per gli scambi culturali. Con l'abbattimento delle distanze geografiche, reso possibile nell'ultimo secolo dai mezzi di trasporto e ancor più dai mass media, l'atavico desiderio di entrare in relazione con l'alterità ha assunto un respiro globale, sostenuto e alimentato dalla grande facilità di creare contatti e raggiungere un numero crescente di individui. Le nuove opportunità offerte hanno trovato spazio e applicazione in ogni ambito. A giovare dell'apertura al mercato globale è stata in primo luogo l'economia, ma anche la cultura e le scienze: la possibilità di scambiare conoscenze, altrettanto se non ancor più importante dello scambio di beni e merci, ha consentito il grande avanzamento scientifico e tecnologico che ha cambiato il mondo e la vita di ciascuno. Non solo le scienze classiche si sono evolute ma ne sono sorte di nuove e con un livello elevato di specializzazione e tecnicità. Non rappresenta quindi un dato sorprendente l'aumento esponenziale, nello stesso periodo, della produzione di documenti scritti e pubblicati, maggiore della produzione totale di tutti i tempi. Austerhmühl (2001: 6-7) registra l'esistenza di circa 170.000 riviste scientifiche attualmente pubblicate e la produzione media di 20.000 nuovi articoli scientifici al giorno. Tali dati sono ad oggi sicuramente da rivedere al rialzo, e ad essi deve essere aggiunta la produzione di documenti di natura commerciale, letteraria, legale, giornalistica, la vasta quantità di testi pubblicati ogni giorno sul web ecc. Conseguentemente a ciò, si è anche registrata, a partire dagli anni '50, una crescita molto importante della richiesta di traduzioni, che ha determinato la necessità di rivoluzionare e aggiornare questa attività vecchia come il mondo. Il dover gestire quantità tanto elevate

di testi da tradurre, tipologie testuali tanto diverse, per innumerevoli combinazioni linguistiche e nel rispetto delle esigenze fondamentali del mercato, ovvero qualità, rapidità ed economicità, tutto ciò ha reso necessaria una trasformazione radicale dell'attività di traduzione e una preparazione adeguata dei traduttori alle nuove sfide che avrebbero dovuto affrontare.

Un fattore determinante in tale rivoluzione è stato indubbiamente rappresentato dalla tecnologia che, come si vedrà, fin dal suo primo sviluppo è stata applicata alla sperimentazione linguistica. In particolare, due eventi hanno rappresentato un vero punto di svolta: la diffusione dei personal computer a partire dagli anni '90, e l'avvento di Internet, circa dieci anni dopo (Giammarresi, in Di Sparti, 2006: 114). I PC hanno cambiato completamente il modo di compiere praticamente ogni lavoro, consentendo la gestione di contenuti digitali e l'automazione di alcune attività; i programmi di videoscrittura e la possibilità di salvare documenti in formato elettronico hanno permesso di intervenire in maniera molto più semplice sui testi per modificarli, visualizzare documenti senza il bisogno di stamparli e salvare in memoria risorse anche molto ingombranti nello spazio di un floppy disk o di un CD. La diffusione delle connessioni ad Internet ha rappresentato un ulteriore e fondamentale passaggio. Il World-Wide Web ha aperto un infinito potenziale di scambi e condivisione di documenti, strumenti e risorse, comunicazioni e contatti in tempo reale e senza confini e l'accesso ad una clientela mondiale; ha determinato la nascita di fenomeni quali la globalizzazione e la localizzazione (di cui si parlerà nei prossimi capitoli) e ha cambiato ulteriormente e ancor più radicalmente il modo di percepire il mondo, il lavoro, gli scambi, l'economia (Pym, 2006: 746). L'avvento di Internet ha rappresentato un immenso beneficio ma anche una sfida per il settore della traduzione, che ha dovuto rispondere prontamente ai grandi cambiamenti da esso prodotti e alle nuove esigenze di comunicazione globale. Per far fronte al crescente volume di richieste e al bisogno di offrire servizi competitivi in termini di costo e tempi di realizzazione, iniziarono a diffondersi le agenzie di traduzione o *Language Services Providers*, che offrono servizi linguistici potendo contare su una rete di traduttori freelance dislocati in tutto il mondo e messi in contatto tramite la condivisione di piattaforme web. Ciò presuppone un'organizzazione nuova del lavoro, con la nascita di figure specifiche per le varie fasi di gestione del testo da tradurre e della traduzione stessa. Anche per il traduttore cambia il modo di lavorare. Non ha più un

rapporto diretto col cliente ma con il Project Manager, deve interfacciarsi di continuo con i colleghi per assicurarsi di adottare le stesse scelte linguistiche e stilistiche, deve condividere i materiali consultati o utilizzare quelli forniti dall'agenzia. Ciò da un lato limita la sua autonomia ma dall'altro gli dà la garanzia di stare utilizzando la terminologia concordata con il cliente finale e condivisa da tutto il team. Il luogo fisico di lavoro si sposta dalla scrivania alla rete, che consente un'interazione in tempo reale utile a monitorare ogni aspetto e ogni fase del processo. Per rendere possibile un'organizzazione tanto articolata, la tecnologia è venuta in soccorso con strumenti specifici ed utili risorse, ormai divenuti indispensabili e parte integrante a pieno titolo del bagaglio professionale dei traduttori. Archer (2002, cit. in Quah 2006: 19) sottolinea come i traduttori professionisti debbano possedere oggi *“professional and linguistic skills in connection with the increasing use of new technologies”*. Gli fanno eco Pym (2003 e 2013), Mossop (2006), Somers (2003) e numerosi altri studiosi ribadendo la necessità di integrare le competenze tecnologiche all'interno dell'ampio bagaglio di conoscenze che un traduttore deve possedere.

### **1.1.2. Gli esordi**

I sussidi digitali per la traduzione, altrimenti noti come CAT Tools, comprendono molteplici tipologie di strumenti e risorse digitali atti ad agevolare e sveltire il lavoro del traduttore al fine di accrescere il ritmo di produzione, consentendogli al contempo di monitorare costantemente i parametri qualitativi del suo lavoro, come la correttezza terminologica, la coerenza interna al testo, l'attinenza al contesto linguistico. Essi permettono inoltre la gestione ordinata e puntuale dei progetti, dall'acquisizione del testo fino alla consegna della traduzione al cliente. La nascita e diffusione dei CAT Tools è avvenuta di pari passo con la rivoluzione tecnologica iniziata dalla seconda metà del secolo scorso, con la comparsa dei primi computer e strumenti di elaborazione di testi, che rappresentarono una epocale innovazione nel modo di reperire, creare, consultare, salvare e modificare molto più facilmente documenti in formato testuale. Rudimentali programmi di traduzione fecero la prima comparsa durante il periodo della guerra fredda, quando l'urgenza di decrittare e tradurre messaggi e comunicazioni raccolte dall'intelligence statunitense e sovietica catalizzò l'attenzione delle istituzioni e lo stanziamento di ingenti finanziamenti nello sviluppo dei primi software di traduzione automatica. Numerosi furono i progetti di ricerca avviati da centri universitari ed enti

governativi e militari, che diedero come frutto i primi sistemi di gestione terminologica bilingui o multilingui ed i primi programmi di traduzione, che ambivano alla totale automazione del processo (Giammarresi, in Di Sparti, 2006: 115). Tale impresa tuttavia si rivelò presto molto ardua, oltre che per la giovanissima età degli studi nel campo, anche per la tecnologia disponibile all'epoca, ancora del tutto insufficiente per un progetto tanto ambizioso come la traduzione interamente automatica. I primi calcolatori elettronici, altamente costosi e capaci di elaborare una quantità limitata di dati, erano esclusivamente alla portata di quelle istituzioni che potevano permettersi un investimento tanto cospicuo, ma i risultati poco soddisfacenti che essi producevano fecero presto desistere i governi dei vari Paesi dal continuare a finanziare le sperimentazioni in corso. Malgrado ciò, l'idea di utilizzare i computer per tradurre non perse di attrattiva, si svilupparono i primi studi teorici sul *Natural Language Processing*<sup>1</sup> e nuovi tentativi di automazione del processo traduttivo vennero compiuti col passare degli anni, fino ad oggi, con risultati mai all'altezza delle aspettative, ma sempre migliori e concretamente utilizzabili in alcune limitate circostanze<sup>2</sup>. Attorno agli anni '60-'70, tuttavia, venne avanzato un nuovo approccio sperimentale secondo il quale la macchina dovesse non sostituire, ma assistere il traduttore e facilitarne il lavoro. L'originario e ardito obiettivo di una traduzione totalmente automatica e di ottima qualità (FAHQT – *Fully Automatic High Quality Translation*) lasciò spazio a ipotesi più realizzabili di integrazione tra le competenze umane e le potenzialità dei computer. L'apporto della macchina passò quindi da una – impercorribile per quei tempi – ipotesi di sostituzione totale della traduzione umana, ad una più utile ed efficace proposta di affiancamento e sostegno, articolata in un continuum che prevedeva diversi livelli di meccanizzazione e intervento umano, come mostra lo schema di Hutchins e Somers (1992) sotto riportato. Ai due estremi troviamo da un lato la traduzione umana e dall'altro la traduzione automatica, mentre le posizioni intermedie sono occupate dalle nuove proposte di traduzione umana assistita da computer (MAHT) e traduzione automatica sostenuta dal traduttore umano. Tali opzioni, che vedono una preponderanza ora della macchina, ora del traduttore, sono la base da cui si è sviluppato il concetto di CAT Tool.

---

<sup>1</sup> Per una esauriente introduzione sull'argomento, si veda Liddy, 2001

<sup>2</sup> Per una trattazione più estesa sulla nascita dei sistemi di traduzione automatica si rimanda al Cap. 5, Seconda Parte del presente lavoro.

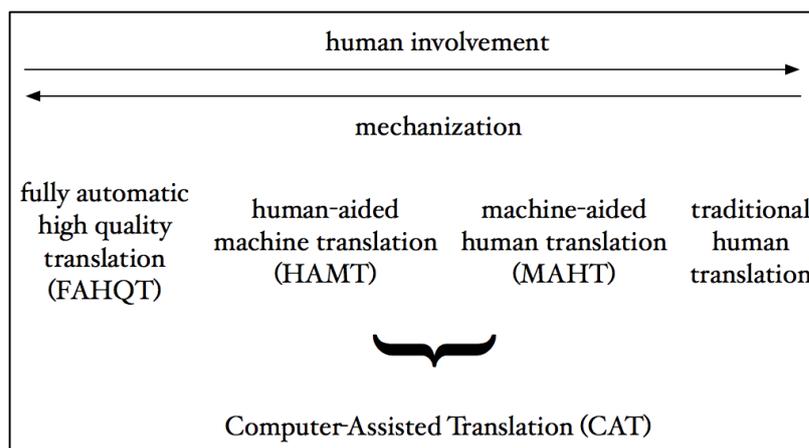


Fig. 1 – Livelli di automazione in traduzione (Hutchins e Somers, 1992)

Tra i pionieri di questo nuovo approccio vanno annoverati Yehoshua Bar-Hillel, ricercatore del MIT, che propose per primo l'introduzione dell'intervento umano all'inizio o alla fine del processo di traduzione automatica, rinunciando quindi alla sua completa automazione (Nirenburg, 1996), e Martin Kay, che nell'introduzione al saggio "The Proper Place of Man and Machines in Language Translation" (1980), così esprime il condiviso entusiasmo per il contributo dei computer nei confronti delle nuove e più grandi esigenze di traduzione:

*"The world is badly in need of translators. (...). But we are fortunate to be children of the age of computers and it is to them that we naturally turn. A computer is a device that can be used to magnify human productivity. Properly used, it does not dehumanize by imposing its own Orwellian stamp on the products of the human spirit and the dignity of human labor but, by taking over what is mechanical and routine, it frees human beings for what is essentially human. Translation is a fine and exacting art, but there is much about it that is mechanical and routine and, if this were given over to a machine, the productivity of the translator would not only be magnified but his work would become more rewarding, more exciting, more human."* (Kay, 1980: 1)

L'intenzione, quindi, non è quella di separare il lavoro di macchina e uomo, quanto piuttosto quello di fare in modo che l'uno sostenga l'altro al fine di migliorare il prodotto

in termini sia qualitativi che quantitativi, e al fine di rendere il lavoro del traduttore più professionale, più appagante e meno meccanico e ripetitivo.

### **1.1.3. I primi CAT Tools**

CAT è un acronimo che sta per *Computer-Aided* (o *Assisted*) *Translation*, quindi i CAT Tools sono letteralmente strumenti computerizzati di sussidio alla traduzione. I primi strumenti CAT furono prodotti alla fine degli anni '60 a partire dalle ricerche condotte nell'ambito della linguistica computazionale e della traduzione automatica. Si trattava essenzialmente di sistemi capaci di reperire informazioni da *database* terminologici o raccolte digitali di testi allineati (testo fonte e traduzione) secondo il principio dei corpora bilingui. Il loro scopo principale era quello di velocizzare la ricerca della terminologia e di individuare la presenza di parti di testo già precedentemente tradotte all'interno dei documenti di riferimento e riutilizzabili per la nuova traduzione. Da questi prototipi, tra la fine degli anni '70 e i primi anni '80, furono ideate e sviluppate le prime TM (Translation Memories), che rappresentano oggi la base di tutti i CAT Tools (Fernandez – Rodriguez, 2010). Tra i primi sistemi si possono annoverare TSS (Translation Support System), prodotto da Alpnet, e TextTools della olandese INK (Reinke, 2013: 29). Si trattava di software contenenti un elaboratore di testi multilingue ed un sistema di gestione della terminologia, e capaci di salvare e recuperare frasi precedentemente tradotte. Contenevano inoltre dizionari bilingue espandibili e funzioni di analisi dei testi e miglioramento della tastiera con l'aggiunta di caratteri speciali. Text Tool fu inoltre lo strumento a partire dal quale una società di traduzioni operante a Stoccarda per IBM negli anni '80 divenne la prima produttrice di CAT Tools a livello mondiale: Trados, società leader nel settore, alla cui suite di prodotti si farà riferimento nei prossimi paragrafi.

Il costo di tali sistemi, tuttavia, era ancora molto elevato e alla portata solo delle grandi società, e bisognò aspettare gli anni '90 per le prime versioni commerciali o open-source dei software per la traduzione, accessibili anche alle piccole aziende e ai traduttori freelance. Nell'arco degli ultimi 25 anni i CAT Tools sono diventati degli strumenti indispensabili che ogni traduttore freelance deve saper padroneggiare e dimostrare di conoscere per poter essere competitivo sul mercato.

## *1.2 L'impatto della tecnologia sul lavoro del traduttore*

### **1.2.1. Tecnologia e traduzione**

Come già affermato precedentemente, la tecnologia ha cambiato il modo di svolgere ogni lavoro e di pensare al mercato e alle relazioni interpersonali e commerciali (Folaron 2010a). I continui progressi della scienza e della tecnica impongono di mantenersi al passo coi tempi per non restare indietro in un mercato in cui la competitività è divenuta molto elevata. L'avvento dei computer e la loro affermazione sempre più indiscussa, insieme alla diffusione di internet, non hanno modificato direttamente il processo principale di traduzione inteso come passaggio di un contenuto da una lingua a un'altra, ma hanno sicuramente rivoluzionato nella pratica il modo in cui esso è compiuto, in termini di organizzazione e procedure, ricerca e consultazione dei materiali di riferimento, analisi del testo, ecc. A livello più profondo, come si vedrà nel Terzo capitolo della presente tesi, è possibile affermare che l'applicazione degli strumenti digitali all'interno del processo traduttivo, dall'acquisizione del testo alla consegna della traduzione, ha avuto un impatto anche sui processi mentali consapevoli e inconsapevoli messi in atto dal traduttore. In questo momento tuttavia ci si soffermerà ad analizzare soltanto l'impatto che la tecnologia ha avuto a livello pratico sul lavoro del traduttore.

Disparti (2004: 40-41) così sintetizza le necessità alle quali la tecnologia può dare risposta durante il processo di traduzione:

- a) Evitare la ripetizione di attività già svolte, come la traduzione di parti di testo già tradotte in passato (si pensi ai casi delle versioni aggiornate di manuali tecnici o cataloghi contenenti la descrizione di articoli commerciali);
- b) Creare forme di collaborazione su LAN o su Internet con i traduttori coinvolti nel processo;
- c) Gestire le grandi quantità di documenti da tradurre;
- d) Riuscire a organizzare i tempi ristretti di consegna;
- e) Rispettare gli standard terminologici fissati dal cliente o dettati da standard internazionali;
- f) Ricorrere ad un unico strumento che funga da ambiente virtuale di traduzione in cui svolgere tutte le funzioni richieste.

### 1.2.2 Esigenze comunicative e organizzative

In primo luogo, l'opportunità di comunicare su scala mondiale e in tempo reale ha aperto nuovi orizzonti e prospettive di mercato, facendo sì che la traduzione divenisse un business internazionale a pieno titolo, con società che servono clienti di paesi diversi, collaborano con traduttori con sede in tutto il mondo e offrono una varietà di servizi linguistici a misura delle esigenze della clientela. Quindi, il primo contributo che possiamo annoverare è proprio quello offerto dai mezzi di comunicazione (Biau Gil e Pym, 2006: 6). Le email in particolar modo hanno rappresentato un sostituto molto più comodo e funzionale di telefono e fax, permettendo comunicazioni istantanee tra le diverse figure coinvolte nel processo e l'invio agevole di documenti in formato digitale. Oltre alla posta elettronica la tecnologia mette a disposizione dei traduttori e dei *project manager* numerosi altri strumenti per la comunicazione, come i programmi di messaggistica istantanea (Skype, Messenger, Lync ecc.) per le comunicazioni urgenti, le Wiki e i forum online per la condivisione di materiali e la discussione, il confronto e l'assistenza tra colleghi, i programmi per audio o videoconferenze per sessioni di formazione o discussioni che richiedono l'interazione diretta tra le parti. Alcuni di questi strumenti verranno descritti più in dettaglio nel corso del secondo capitolo. Dei vantaggi dei documenti in formato digitale si è già fatto accenno al paragrafo precedente: essi consentono di consultarne il contenuto e modificarlo se necessario, senza dover stampare alcunché e mantenendo inalterate quelle parti del documento che devono essere mantenute tali, come per esempio la formattazione, il layout, eventuali immagini ecc. Inoltre possono essere salvati nella memoria del computer o su supporti hardware, ordinati in cartelle e sottocartelle che ne rendano rapido e semplice il reperimento (Pym, 2006: 745). Lavorare su documenti digitali richiede tuttavia oltre al possesso di diversi programmi per poter visualizzare e modificare file in formati diversi, anche le competenze necessarie a lavorare su tali formati, conoscerne le interfacce grafiche, sapere dove reperire i comandi ecc. In alcuni casi implica anche il dover padroneggiare alcune conoscenze più specifiche, come i linguaggi di markup (HTML o RTF), o programmi non tipicamente utilizzati per uso linguistico, come ad esempio i CAD. Come si vedrà, i CAT Tools vengono in aiuto in questo senso convertendo diversi tipi di file in un formato specifico che consente di modificarli con facilità, per poi riconvertirli nel formato originario. Tuttavia, appare chiaro che un traduttore deve essere pronto ad equipaggiarsi

con un'adeguata conoscenza degli strumenti tecnologici, inserendo nel suo percorso formativo dei corsi appositi o approfondendo l'argomento autonomamente.

### **1.2.3. I tempi di lavoro e le agenzie**

Forse il vantaggio più conclamato offerto dal ricorso agli strumenti tecnologici in traduzione è la maggiore rapidità dei tempi di lavoro, in risposta alla crescente richiesta e all'esigenza di competitività del mercato (Folaron, 2010): maggiore la velocità di lavoro, maggiore il quantitativo di traduzioni che si riesce a produrre, con conseguenti maggiori profitti. Grazie alla consultazione di glossari e dizionari in formato digitale, all'utilizzo delle porzioni di testo già tradotte e salvate in memoria e alla possibilità di far svolgere al computer alcune attività di verifica e monitoraggio della qualità del lavoro che richiederebbero altrimenti considerevole tempo, attenzione ed energie da parte dei traduttori, i ritmi di lavoro divengono più snelli e la produttività maggiore. Relativamente alla qualità, essa è tutelata non solo dalla competenza del traduttore, ma anche dalle varie fasi di analisi e *quality check* che avvengono sia con gli strumenti preposti, sia con il lavoro di *editor* e *proofreader* che rivedono il lavoro dei traduttori durante diverse fasi di revisione e rilettura, al fine di garantire quanto più possibile un risultato di elevata qualità.

La traduzione all'interno delle agenzie è un processo complesso, che richiede capacità gestionali oltre che linguistiche. In diversi casi le esigenze di comunicazione globale rendono necessario che un documento venga tradotto in più lingue contemporaneamente<sup>3</sup>, per cui il lavoro di gestione dei progetti diventa più articolato e necessita di strumenti specifici che consentano di svolgere tutti i compiti di *project management* in maniera ordinata e funzionale: ricevere il testo, dividerlo in parti da inviare alle decine di traduttori dislocati in diversi paesi, creare, condividere e aggiornare linee guida e glossari multilingue, inviare e ricevere dai collaboratori i testi da tradurre e tradotti, ecc. Poter contare su supporti specifici per svolgere queste azioni in maniera attenta e puntuale è essenziale, tanto quanto condividere gli stessi strumenti con tutto il team di collaboratori coinvolti nel progetto. Proprio per questa esigenza, molte agenzie di traduzione internazionali preferiscono dotarsi di un portale internet proprio, tramite il quale gestire i progetti, creare e condividere le memorie di traduzione, far tradurre i testi e salvarli in memoria, mantenere aggiornati i glossari ed eventuale altro materiale di

---

<sup>3</sup> Cf. par. seguente sulla localizzazione

supporto in modo che essi possano essere agevolmente utilizzati, modificati e resi disponibili in tempo reale a tutto il team.

I vantaggi dei singoli strumenti verranno più dettagliatamente descritti nel prossimo capitolo (cf. par. 2.1, 2.2. e 2.3). Occorre in questo momento anche evidenziare che alcuni traduttori, soprattutto quelli che hanno lavorato per anni in maniera “tradizionale”, sono restii all’adozione di strumenti CAT per via dell’investimento non solo economico, ma anche in termini di tempo, che comporta l’apprendimento dell’uso del software e l’acquisizione di una nuova prassi di lavoro. Come infatti sottolinea Lagoudaki (2006), i benefici derivanti dall’adozione di un CAT Tool sono visibili principalmente nel lungo periodo, mentre all’inizio la loro introduzione implica una fase di apprendimento e una ridiscussione delle proprie abitudini e dei propri metodi di lavoro che può rallentare i ritmi e ridurre la produttività. Inoltre, come tutti i traduttori che lavorano con i CAT Tools sanno, questi strumenti si rivelano più o meno utili a seconda dell’ambito linguistico per il quale si traduce, quindi occorre saperli impiegare con spirito critico e capacità di discernimento. Tuttavia, una volta vinte queste iniziali difficoltà, si potrà contare su di un supporto di grande utilità e versatilità. Oggi sempre più corsi universitari di traduzione includono al loro interno laboratori di CAT Tools (cf. cap. 4), quindi i nuovi traduttori sono maggiormente consapevoli, rispetto ai loro colleghi più anziani, dei benefici di tali strumenti e dell’importanza di conoscerne e padroneggiarne l’utilizzo per potersi affermare professionalmente sul mercato, in cui questo genere di competenze viene richiesto con condizione indispensabile, al fianco di una solida formazione linguistica.

### *1.3 L’industria della traduzione<sup>4</sup>*

Quando si parla di industria della traduzione, si fa riferimento a un mercato strutturato, che come qualsiasi altro settore del commercio o della prestazione di servizi implica l’incontro tra domanda e offerta, e un’organizzazione in cui professionisti diversi collaborano alla realizzazione di un prodotto finale. Ai cultori della traduzione letteraria e puramente filologica, tale definizione potrebbe risultare indubbiamente ostica e persino aliena. Tuttavia, come già accennato, la traduzione commerciale è un fenomeno in continua espansione che poco ha a che vedere con gli studi teorici e linguistici, o con

---

<sup>4</sup> Molto interessante, sulle dinamiche traduttore – agenzia a cui si farà qui riferimento, il saggio di Olohan (2011)

l'approccio comunicativo o cognitivo alla traduzione. Questo fenomeno guarda alla traduzione da una prospettiva nuova e molto più concreta, come un'esigenza fondamentalmente economica e pratica, come un prodotto o un servizio alla pari di tanti altri prodotti e servizi offerti sul mercato.

### **1.3.1. Una descrizione**

Olvera *et al.* (2005: 133 – 135) descrivono il mercato della traduzione con pochi ma eloquenti aggettivi. È un mercato *globale* perché non contempla confini, allo stesso modo dei commerci e della diffusione delle informazioni al giorno d'oggi. È *decentralizzato* e *virtuale* perché la distanza fisica, grazie alle nuove tecnologie, non è più un problema per l'assegnazione di incarichi di traduzione che rispondano all'esigenza di economicità e qualità. Internet diventa il luogo degli scambi e dei contatti tra agenzie, professionisti e clienti. È un mercato *specializzato* che deve rispondere a infinite esigenze, sia in termini di combinazioni linguistiche che di settorialità e specificità degli ambiti linguistici richiesti e delle finalità comunicative perseguite. È una realtà *dinamica*, in continua espansione con l'affermarsi di settori d'interesse sempre nuovi e specifici, che richiedono ai traduttori una continua formazione. È infine un mercato *esigente*, che richiede ai traduttori preparazione, competitività, versatilità, tempi ristretti e il rispetto di regole diverse a seconda del cliente o del prodotto. Gli operatori di tale industria sono essenzialmente le agenzie di traduzione o *Language Services Providers* (LSP), che possono avere una o più sedi operative ufficiali, o possono svolgere il loro lavoro interamente tramite il Web. All'interno delle agenzie operano diverse figure specializzate, che possono essere dipendenti *in-house*, o collaboratori autonomi a distanza, come i *freelancer*. Le maggiori agenzie di traduzione possono contare sia su personale *in-house* che su una fitta rete di collaboratori *freelance*, con mansioni diverse. Internamente vengono gestiti solitamente i rapporti con i clienti, la contabilità, la manutenzione dei server contenenti le centinaia di memorie di traduzione, glossari e corpora, e la fase di *project management* dalla ricezione dei testi da parte del cliente alla consegna finale delle traduzioni. Le figure professionali in ambito linguistico che lavorano in agenzia sono quindi i project manager e gli addetti alla cura delle TM e della terminologia, con la manutenzione dei glossari e dei materiali di consultazione. Le mansioni che vengono invece spesso affidate a collaboratori *freelance* sono la traduzione e la revisione. Le combinazioni linguistiche gestite da un'agenzia sono potenzialmente infinite, così come

gli ambiti linguistici di specializzazione. Il numero di collaboratori necessari a garantire tale varietà di servizi è quindi molto elevato, e non gestibile sia fattivamente che economicamente con personale dipendente. Traduttori, editor e revisori sono quindi selezionati da un ampio database di collaboratori, in base alle loro combinazioni linguistiche e ambiti di specializzazione, e retribuiti in funzione della quantità di lavoro svolto, calcolato con tariffe a parola, a ora o a cartella, a seconda della comune convenienza. Il lavoro in team è prassi comune, specialmente quando si devono gestire progetti molto voluminosi in tempi brevi. Il coordinamento del team di traduttori e revisori è affidato al project manager, che dividerà i materiali da tradurre, condividerà le risorse linguistiche da consultare durante il lavoro e si accerterà che ogni membro del team rispetti le consegne.

Sono diversi i servizi linguistici offerti dalle LSP. In primo luogo, ovviamente, servizi di traduzione e di interpretariato, quest'ultimo in presenza o anche tramite telefono o programmi di videoconferenza. Oltre a questi due principali servizi, vengono sempre più richieste e offerte opere di trascrizione da audio, revisione di testi, sottotitolaggio di video e loro eventuale traduzione, ecc.

### **1.3.2. Il fenomeno della localizzazione**

Un particolare servizio offerto spesso da agenzie specializzate è quello della localizzazione, la cui affermazione a partire dagli anni '90 ha contribuito significativamente all'affermazione del settore dei servizi linguistici sul mercato globale. Il termine *localizzazione*, o per meglio dire il suo equivalente inglese *localization*, viene abbreviato con *L10n*, dove il 10 rappresenta il numero di lettere tra la *L* e la *n*. Esso indica una traduzione che è anche "adattamento culturale", come suggerisce una delle sue definizioni ufficiali, che recita: "*Localization involves taking a product and making it linguistically and culturally appropriate to the target locale (country/region and language) where it will be used and sold*" (LISA, cit. in Esselink 2003: 67) In un mercato senza confini, sia i prodotti commercializzati che la documentazione che li accompagna spesso necessitano non solo di essere tradotti, ma di essere anche specificamente adattati alle normative commerciali del paese o semplicemente alle aspettative culturali o alle prassi in esso in uso. Come sottolinea la definizione sopra riportata, il prodotto deve essere *linguisticamente* ma anche *tecnicamente* appropriato, modificandone se del caso le specifiche per renderle coerenti con gli standard del mercato di destinazione. Nel

tempo, il termine ha assunto una connotazione più specificamente legata all'ambito tecnologico, principalmente alla localizzazione dei software e, più recentemente, dei siti Web e dei videogiochi. Questo particolare campo di specializzazione necessita sia di adeguate competenze linguistiche ai fini della traduzione, ma anche di un insieme di competenze tecniche molto specifiche, in quanto i testi da tradurre sono l'anima stessa del prodotto, fanno parte dei suoi meccanismi interni e ne determinano il corretto funzionamento: il software manipola i testi sotto forma di stringhe di byte, li memorizza e li elabora. I comandi vengono dati inserendo o selezionando del testo, altro testo spiega interattivamente all'utente il funzionamento del programma e lo avvisa di eventuali condizioni di errore. Qualsiasi errore tecnico nella traduzione di tali materiali può portare a malfunzionamenti del programma, per questo il traduttore-localizzatore sarà responsabile non solo della correttezza della traduzione, ma anche del corretto funzionamento del prodotto localizzato. È necessario quindi che il traduttore che si occupa di localizzazione acquisisca specifiche competenze tecniche. Dovrà padroneggiare gli standard terminologici e stilistici, e al contempo destreggiarsi agevolmente fra strumenti informatici e aspetti tecnici come la conoscenza di formati di file, delle basi di programmazione dei software ecc. I localizzatori devono inoltre investire in attrezzature hardware e programmi software aggiornati, oltre che nei comuni programmi per ufficio. Sarà indispensabile un software di traduzione assistita, in grado di gestire e convertire diversi formati di file e database testuali, una connessione internet ad alta velocità e permanente per la gestione dei progetti di localizzazione di siti web, nonché la disponibilità di versioni installate dei più diffusi sistemi operativi, sulle quali verificare le procedure descritte e i dubbi terminologici che i glossari non sono in grado di risolvere. I traduttori – localizzatori sono stati i primi ad avvalersi dei CAT Tools, poiché sia le caratteristiche del materiale da tradurre (elevata ripetitività, grandi volumi, argomento tecnico) che l'ambiente di lavoro (fortemente computerizzato e tecnologico) si sono ben prestati all'introduzione di strumenti quali i database terminologici e le memorie di traduzione.

### **1.3.3. Localizzazione vs. traduzione**

L'affermarsi del fenomeno della localizzazione è stato accompagnato da una discussione sulla sua definizione, soprattutto in rapporto alla traduzione. Veniva infatti sottolineato quanto queste due attività avessero in comune, e come in fondo entrambe

riguardassero la trasposizione di contenuti da una lingua a un'altra. In realtà, la localizzazione è più di una mera traduzione interlinguistica. Come mostrano gli schemi sotto riportati (Esselink 2000: 10 e 14), la localizzazione è un processo articolato che comprende diverse fasi e coinvolge diversi attori. Alcune di queste fasi e di questi attori sono comuni a quelli chiamati in gioco dall'attività di traduzione, altri invece sono peculiari, per cui si può affermare che la traduzione è solamente una delle componenti di un progetto di localizzazione:

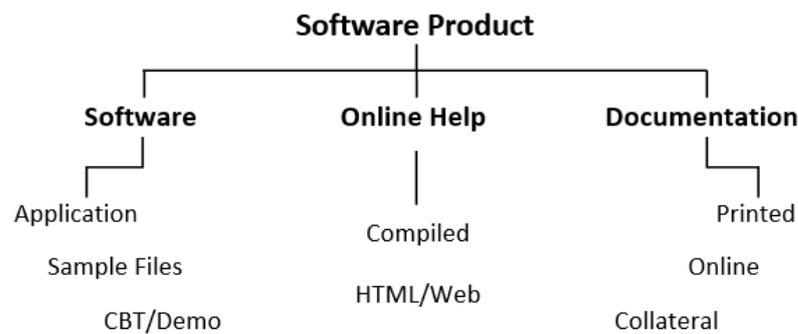


Fig. 2: Le componenti del progetto di localizzazione (Esselink 2000: 10)

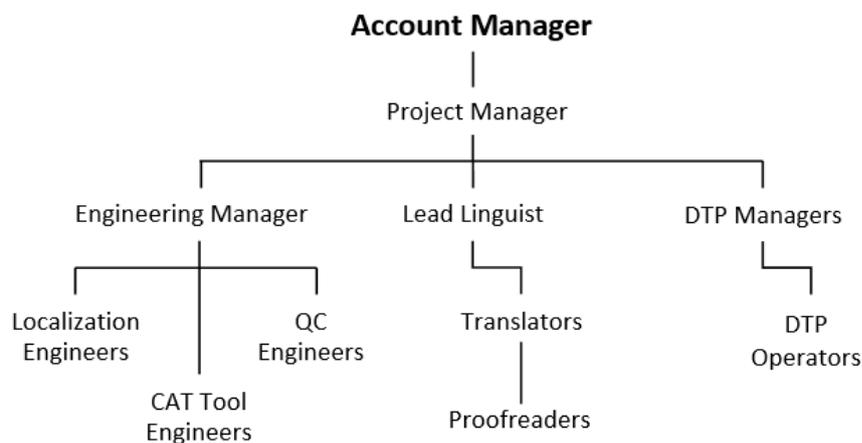


Fig. 3: I professionisti coinvolti nel progetto di localizzazione (Esselink 2000: 14)

La localizzazione è un'attività molto varia, può riguardare applicazioni molto diverse, da sistemi relativamente statici e distribuiti su grande scala come dei database, a contenuti in continua evoluzione come le pagine web. È inoltre un'attività molto dinamica e dalle tempistiche ristrette, data la necessità di lanciare nuovi software o nuovi contenuti web contemporaneamente in più lingue e per più mercati (la cosiddetta "simship"). Per accelerare le tempistiche e garantire che il processo di localizzazione avvenga senza difficoltà potendo così offrire un prodotto perfettamente funzionale per i mercati

internazionali, già in fase di progettazione e sviluppo del software i programmatori possono provvedere alla sua internazionalizzazione (*I18n*). Con questo termine si intende la preparazione del contenuto digitale ad un suo uso in lingue diverse e alla sua localizzazione, ad esempio tramite la predisposizione del software per i caratteri *double-byte* necessari per visualizzare gli ideogrammi delle lingue asiatiche, la capacità di supportare i set di caratteri speciali come le vocali accentate utilizzate in alcune lingue, la disponibilità di formati diversi per date, orari, numeri e valute, ecc. Gli interventi di internazionalizzazione riducono drasticamente i tempi e i costi della localizzazione, limitando il ricorso a possibili revisioni dei software e alle conseguenti attività di verifica del loro funzionamento. Anche il ricorso a materiali già tradotti e salvati in memorie di traduzione apposite contribuisce a velocizzare, nonché a standardizzare, il processo di traduzione e localizzazione.

#### **1.3.4. Siti web e videogames**

Così come i software, anche le pagine web ed i documenti in formato HTML hanno il bisogno di essere progettate e sviluppate per una loro traduzione in lingue diverse. Il World Wide Web sta diventando sempre più multilingue, ed è il canale preferenziale con cui le società cercano di infiltrarsi nei mercati esteri. Nel caso delle pagine web, il processo di pubblicazione in altre lingue viene denominato “globalizzazione di un sito web”. Un sito ben globalizzato e internazionalizzato limita l’intervento di localizzazione alla semplice traduzione dei contenuti, piuttosto che allo svolgimento di azioni più complesse e costose. Un altro prodotto che viene regolarmente sottoposto a localizzazione, e che ha un indiscutibile potere economico sul mercato, è rappresentato dai videogiochi (De Pedro Ricoy, 2007). Essi hanno da tempo superato i profitti di cinema e programmi televisivi (Chiaro in Munday 2009), e i maggiori produttori di videogiochi, Sony, Nintendo e Microsoft, sono società con una forte presenza internazionale, che si riflette nella promozione e diffusione dei loro prodotti in tutto il mondo. I videogiochi possono essere definiti, nella maniera più ampia possibile, come “*computer-based entertainment software ... involving one or multiple players in a physical or networked environment*” Frasca (2001: 4). Il termine “videogioco” può in effetti inglobare un’ampia varietà di tipologie di giochi, disponibili su supporti diversi (computer, televisore, console, telefoni cellulari), e di difficoltà molto variabile. Anche la complessità dell’interfaccia grafica e dei comandi è molto diversa da gioco a gioco. Essi

mettono insieme immagini, effetti sonori, dialoghi e un'ambientazione che molto spesso riproduce la realtà curandola nei minimi dettagli. In genere, i dialoghi sono sia sonori che trascritti, quindi al momento della produzione delle versioni per altri paesi, i videogiochi vengono sottoposti a traduzione, localizzazione, doppiaggio e sottotitolaggio. È un compito impegnativo per i traduttori, e per le case produttrici un importante passaggio che avviene al momento stesso della realizzazione del prodotto. Nel tradurre e localizzare un videogioco è importante tenere in conto l'effetto finale, che deve garantire il massimo del coinvolgimento e dell'intrattenimento. Nulla deve "stonare" e far perdere al giocatore l'impressione di trovarsi all'interno di uno scenario e una ambientazione reale o fantastica, ma comunque mai approssimativa. L'esempio sotto riportato è un celebre errore di traduzione di un videogioco giapponese del 1991, *Zero Wing*. La frase inglese, evidentemente erronea e inserita in un dialogo in cui molte altre battute sono state tradotte in maniera impropria, è divenuta un tormentone per gli appassionati del genere.



Fig. 4. Una schermata del videogioco Zero Wing, con la celebre frase tradotta erroneamente.

Il traduttore di videogiochi deve tenere presenti fattori culturalmente determinati ma anche le caratteristiche del genere stesso e il tipo specifico di registro che esso utilizza. Per questo motivo gli viene concessa la maggiore libertà possibile nel fare uso di riferimenti locali e culturali, così come di frasi idiomatiche ed espressioni gergali, pur di mantenere il risultato finale e rendere il gioco piacevole e coinvolgente.

## 1.4 Studi linguistici sui CAT Tools

A causa della loro natura principalmente pratica e funzionale, nei primi tempi della loro comparsa gli strumenti digitali per la traduzione non sono stati oggetto di riflessioni e studi teorici strutturati, probabilmente perché la loro applicazione era ancora molto limitata e sperimentale. La nascita di un filone di indagine a livello accademico è quindi relativamente recente, essendo avvenuta contemporaneamente alla diffusione sul mercato di CAT Tools, programmi di traduzione automatica, strumenti di gestione terminologica ecc. (Bowker, 2002).

### 1.4.1. Gli Applied Translation Studies

Lo studio dei CAT Tools e degli altri strumenti di sussidio alla traduzione ha trovato collocazione all'interno della branca degli *Applied Translation Studies*, secondo lo schema di Holmes sotto riportato, presentato per la prima volta alla Terza Conferenza sulla Linguistica Applicata tenutasi a Copenaghen nel 1972<sup>5</sup>:

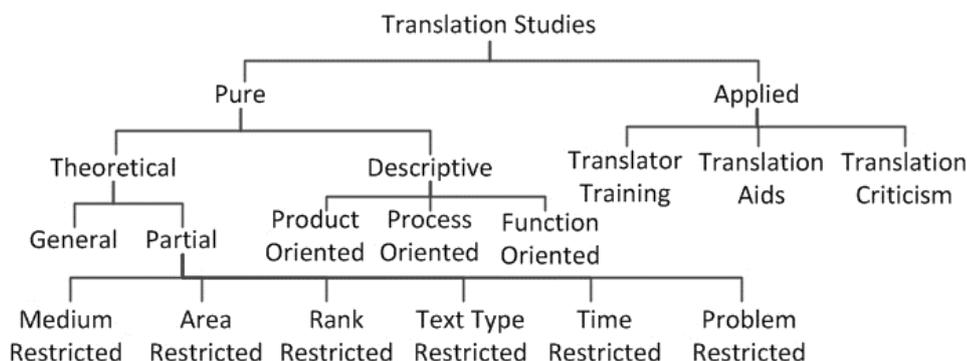


Fig. 5. La divisione dei Translation Studies secondo Holmes (1972)

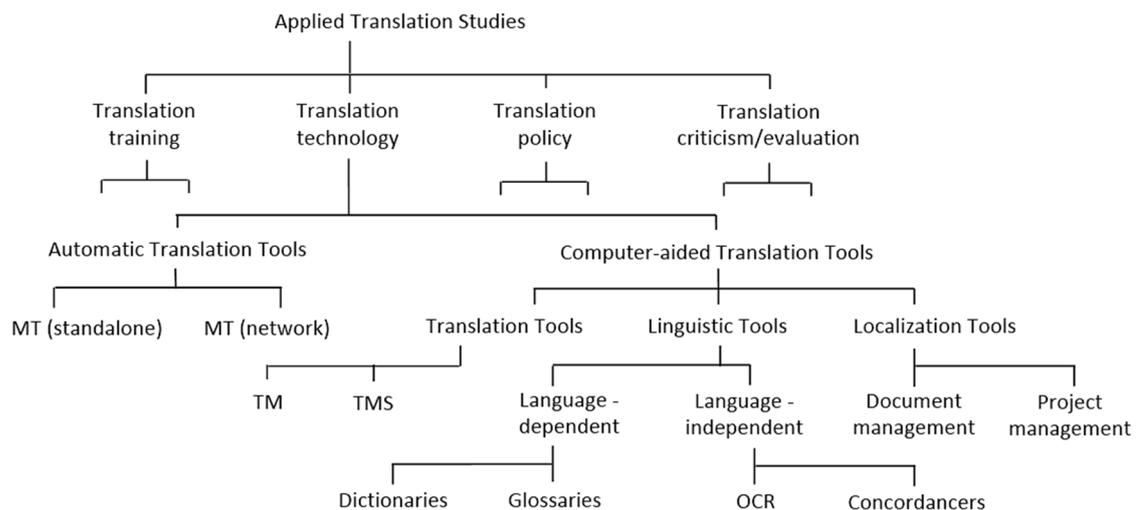
Questo schema utilizza categorie semplici e organizzate gerarchicamente per mostrare la vastità e l'interdisciplinarietà degli studi, che coinvolgono linguistica, sociologia, neuroscienze, psicologia cognitiva, studi sulla cultura, informatica ecc.

I Translation Studies si dividono prima di tutto in due rami, uno detto “puro” (studi teorici e descrittivi), e uno “applicato”. Queste due grandi categorie si dividono a loro volta in sottocategorie. Dallo schema è evidente come il primo gruppo sia molto più articolato rispetto al secondo, segno di come le discipline “pure” avessero avuto una decisa

<sup>5</sup> Pubblicato in Holmes 1975

preponderanza rispetto a quelle applicate fino al momento in cui lo schema è stato creato. Inoltre, lo schema ad albero adottato è unidirezionale, mentre in realtà le categorie individuate sono spesso interconnesse e in alcuni casi sovrapposte le une alle altre. Per esempio, relativamente all'argomento qui trattato, la riflessione sui sussidi digitali alla traduzione ricade sicuramente all'interno della sottocategoria *Translation Aids* degli Applied Translation Studies, ma innegabilmente essa può collegarsi con altre categorie dello schema anche appartenenti ai Translation Studies puri, come *Process-oriented* e *Product-oriented Descriptive TS*, e *Partial Theoretical TS* e tutte le sue sottocategorie. Relativamente agli Applied Translation Studies, le tre sottocategorie identificate da Holmes sono certamente interconnesse le une alle altre, ma possono anche essere ricondotte a studi teorici o descrittivi. Da essi prendono le informazioni e le conoscenze necessarie a sviluppare i propri studi, che a loro volta potranno dare un contributo alle discipline pure, per avanzare nei loro ambiti di ricerca.

Partendo dallo schema di Holmes, e sostituendo la nomenclatura “*Translation Aids*” con “*Translation technology*”, Quah (2006: 42) propone una dettagliata articolazione dei vari tipi di supporti digitali alla traduzione, colmando almeno in parte la mancanza di specificità che caratterizza il lato destro dello schema di Holmes:



MT = Machine Translation; TM = Translation Memory; TMS = Terminology Management Systems;  
OCR = Optical Character Recognition

Fig. 6. Schema degli Applied Translation Studies di Quah (2006)

Questo schema fornisce una panoramica degli strumenti che possono essere annoverati sotto l'etichetta di *Translation Technology*, divisi in prima istanza tra strumenti di traduzione automatica e strumenti di traduzione assistita da computer. A loro volta, questi ultimi sono ulteriormente suddivisi in strumenti per la traduzione (memorie e sistemi di gestione terminologica), strumenti linguistici (dipendenti o indipendenti dal linguaggio, dizionari e glossari nel primo caso, riconoscimento ottico e *concordancer* nel secondo), e infine strumenti per la localizzazione, per la gestione dei documenti e dei progetti. Anche per questo schema vale quanto detto per il precedente, ovvero che le categorie in esso contenute sono spesso interconnesse, soprattutto data la diffusione di strumenti e piattaforme multifunzionali che integrano diverse delle applicazioni sopra citate.

#### **1.4.2. Limitatezza degli studi**

Relativamente alla traduzione automatica, essa ha da subito attirato le fantasie e gli entusiasmi dei ricercatori, per la sua ambizione di riprodurre con la macchina il complesso processo umano di comprensione testuale e riformulazione in un'altra lingua. Tale entusiasmo ha dato vita a un corposo dibattito accademico, con numerose pubblicazioni e sperimentazioni dagli approcci diversificati e modelli e strumenti diversi che verranno trattati in dettaglio più avanti (cf. Parte Seconda). Meno corposa è invece la produzione di studi relativi ai CAT Tools, visti pur sempre con interesse dagli specialisti del settore, ma considerati spesso puri strumenti di supporto il cui compito meccanico di memorizzazione e riproposizione dei segmenti di traduzione poco potesse rivelare del processo traduttivo rispetto ai tentativi di riprodurre per intero (o quasi) tale processo al computer. Relativamente ai CAT Tools si è sviluppata una discussione riguardante la loro applicazione didattica o più semplicemente la loro introduzione nel percorso formativo dei giovani traduttori, considerata la grande richiesta che di essi si fa nel mercato e l'importanza di padroneggiarne l'utilizzo. Oltre a tale approccio didattico, in alcuni casi sono stati studiati anche per il modo in cui essi modificano l'organizzazione del lavoro e i processi pratici e mentali messi in atto nel tradurre (cf. Capitolo Terzo). Come puntualizzato da Mossop (2006), non ci sono molti studi che abbiano cercato di osservare il processo traduttivo dall'interno. La teoria della traduzione è passata dal considerare principalmente il testo come centro del processo al considerare le azioni svolte dalla persona coinvolta nel processo, ovvero il traduttore. Tale cambio di prospettiva è avvenuto solo di recente, così che tutti gli studi empirici sulla traduzione sono stati

condotti “in vitro”, all’interno di università piuttosto che sul campo, e spesso senza il coinvolgimento diretto dei traduttori. I CAT Tools potrebbero dare un contributo nell’indagine sul processo traduttivo, dato che essi lo accompagnano in ogni suo passaggio e sono stati prodotti sulla base dello studio dei comportamenti e dei bisogni ricorrenti dei traduttori in risposta alle situazioni di lavoro. In questo modo, i CAT Tools verrebbero in sostegno dell’approccio che nello schema di Holmes viene denominato “*Process-oriented Translation Studies*”, il cui scopo è appunto quello di investigare i processi mentali del traduttore. La traduzione mediata o assistita da computer apporta dei cambiamenti in tali processi, introducendo nuovi passaggi e modificandone altri. Quah (2006) suggerisce che il processo traduttivo condotto con l’ausilio di strumenti digitali implichi sia una traduzione intralinguistica sia una traduzione vera e propria o interlinguistica, secondo lo schema di seguito riportato (Quah 2006: 43):

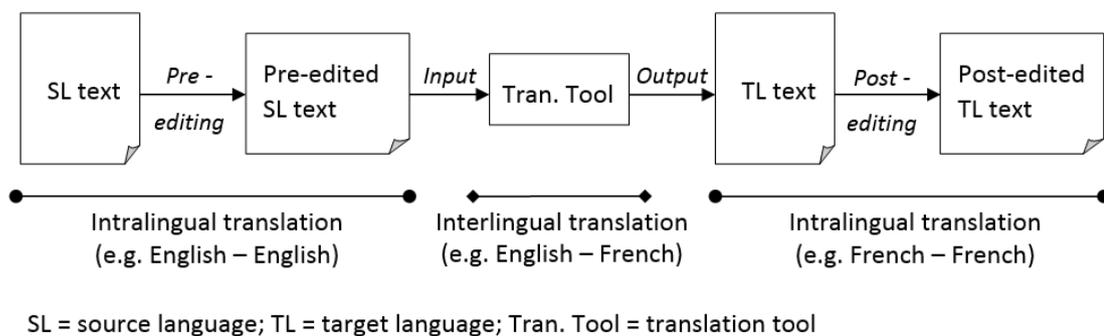


Fig. 7. Un modello del processo traduttivo che include le fasi di pre- e post- editing (Quah 2006)

La traduzione intralinguistica corrisponde quindi alle fasi di pre- e post-editing, tipiche della traduzione sia assistita che prodotta dalla macchina. La fase di pre-editing serve a rendere il testo adatto ad essere elaborato dal software prescelto, e in alcuni casi può implicare una semplificazione del testo al fine di scongiurare possibili ambiguità (utile in particolare per l’utilizzo dei programmi di traduzione automatica, cf. 5.2), o semplicemente un controllo della formattazione del testo per far sì che la segmentazione compiuta dal programma sia stata correttamente eseguita. La fase di post-editing, invece, è decisamente importante dopo la produzione interamente automatica di una traduzione, per correggere gli errori della macchina e garantire che il testo funzioni bene nel complesso. Tuttavia, anche la traduzione umana necessita di un post-editing (o revisione), che può andare dalla semplice verifica di eventuali errori di battitura, grammaticali o

stilistici al controllo dell'uso coerente della terminologia all'interno del testo o del rispetto delle regole stilistiche e formali dettate dal cliente o condivise col team di traduttori.

### **1.4.3. La linguistica computazionale**

Lo studio e la ricerca sui CAT Tools e sugli strumenti digitali per la traduzione in generale si lega anche in maniera particolare ad altri due rami della linguistica, ovvero la linguistica computazionale, o *Language Engineering*, e la linguistica dei corpora. La prima fu resa nota dallo sviluppo dei primi progetti sperimentali di traduzione automatica, ma il suo campo di interesse è molto più ampio, riguardando le diverse forme di interazione tra la macchina e l'uomo attraverso il linguaggio. Si tratta di un campo di ricerca interdisciplinare che mette insieme non solo informatica e linguistica tradizionale ma anche la linguistica cognitiva, la sociolinguistica e discipline scientifiche come la statistica ecc. Il suo scopo è quello di sviluppare intelligenze artificiali per il riconoscimento, la comprensione, l'interpretazione e la produzione di linguaggio umano in tutte le sue forme, e di definire dei modelli a partire dai quali le macchine possano produrre comportamenti linguistici (Chiari, 2008). Essa si centra sul concetto di *Natural Language Processing*, la capacità di un calcolatore elettronico di comprendere ed elaborare un linguaggio naturale per usi e finalità diversi. Secondo Calzolari e Lenci (2004):

*“Attraverso l'analisi linguistica automatica del testo, gli strumenti del TAL<sup>6</sup> sciolgono la tela del linguaggio per estrarre e rendere espliciti quei nuclei di conoscenza che possono soddisfare i bisogni informativi degli utenti. Dotando il computer di capacità avanzate di elaborare il linguaggio e decodificarne i messaggi, diventa così possibile costruire automaticamente rappresentazioni del contenuto dei documenti che permettono di potenziare la ricerca di documenti anche in lingue diverse (Crosslingual Information Retrieval), l'estrazione di informazione rilevante da testi (Information Extraction), l'acquisizione dinamica di nuovi elementi di conoscenza su un certo dominio (Text Mining), la gestione e organizzazione del materiale documentale, migliorando così i*

---

<sup>6</sup> Trattamento automatico della lingua, *nda*.

*processi di elaborazione e condivisione delle conoscenze.*” (Calzolari e Lenci: 2004)

La difficoltà maggiore nel perseguimento di questo obiettivo risiede nella natura stessa del linguaggio, in continua evoluzione e non organizzato secondo poche regole ben definite, quindi difficilmente gestibile da un calcolatore, strumento che nasce essenzialmente per elaborare stringhe di dati assolutamente univoci.

In un primo tempo questa disciplina fu notevolmente influenzata dai metodi deduttivi della grammatica generativa di Chomsky, per cercare di individuare delle regole astratte che descrivessero la competenza linguistica di un parlante, e tentare di applicarle ad una macchina (Chiari, 2008: 60). In un secondo momento fu preferito l'approccio statistico, che valutava tramite un algoritmo il grado di ricorsività di espressioni e costrutti in raccolte di testi di una data lingua per derivarne delle regole stabili. Al fine di rendere questa ricerca più agevole, venne sviluppato un criterio di divisione testuale in porzioni di grandezza variabile, dalle frasi alle parole. Tale processo di divisione del testo prende il nome di chunking o tokenizzazione (Calzolari e Lenci, 2004: 60). Il chunk è la porzione minima di testo che consente di conoscere l'utilizzo reale di una parola in una determinata lingua, grazie alla sua presenza all'interno di un contesto, quindi si tratta di porzioni delle dimensioni di una frase circa. Il token è definito invece come un'unità minima e indivisibile di senso, nella maggior parte dei casi una parola o un'espressione regolare come costrutti e parole composte. La divisione del testo in unità minime è di grande importanza per lo sviluppo dei CAT Tools, sia per le memorie di traduzione che per gli strumenti di ricerca terminologica. La loro capacità di analizzare i testi alla ricerca di materiale già tradotto o terminologia già nota si basa proprio sulla capacità di segmentare i documenti in unità che l'utente può dimensionare a seconda delle sue esigenze, così come avviene in segmenti anche la memorizzazione delle doppiette di testo originario e tradotto.

#### **1.4.4. La linguistica dei corpora**

La linguistica dei corpora ha molto in comune con la linguistica computazionale, che ha fatto ricorso ad essa per l'individuazione delle regole e delle prassi di utilizzo della lingua (Chiari, 2008: 63). La *corpus linguistics* riguarda lo studio del linguaggio a partire dall'osservazione di grandi raccolte digitali di testi di uso reale, i corpora appunto, di

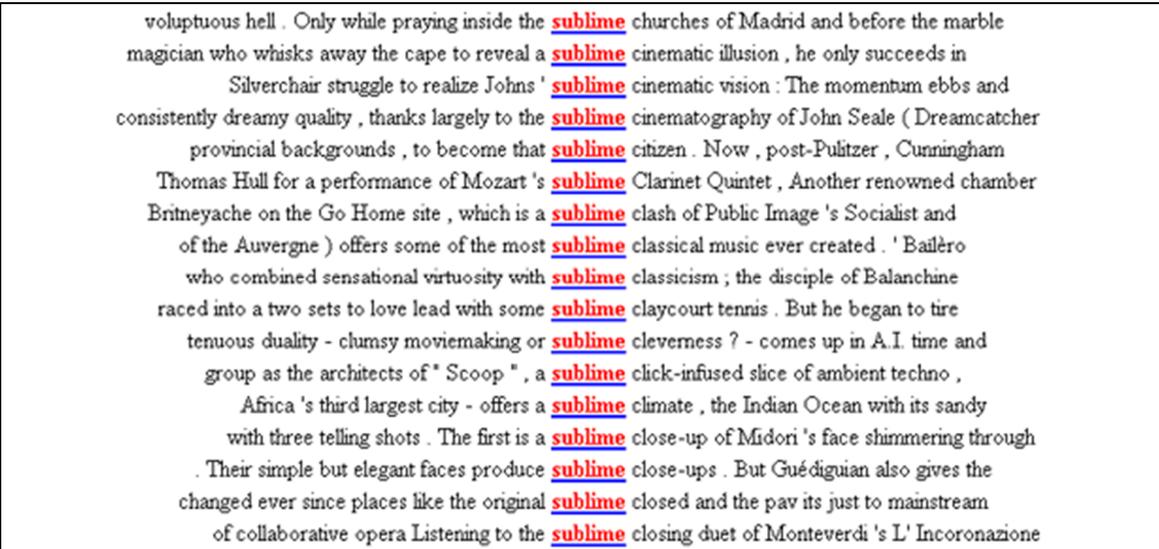
natura generica o ordinate secondo criteri specifici per rappresentare determinati ambiti del linguaggio<sup>7</sup>. I corpora possono inoltre essere monolingui o bi- e plurilingui, per studiare una sola lingua o permettere un confronto tra due o più. I corpora bilingui possono essere semplicemente raccolte di testi in due lingue sullo stesso argomento, o raccolte di testi in una lingua e della loro traduzione in un'altra lingua, allineati secondo una corrispondenza frase per frase. In questo caso essi prendono il nome di corpora paralleli, e sono quanto di più simile ad una memoria di traduzione creata con un CAT Tool. La funzione di allineamento, che si vedrà più in dettaglio nel prossimo capitolo, viene utilizzata anche per la creazione di TM (translation memories) a partire da documenti già tradotti. Lo strumento utilizzato per consultare un corpus viene detto *corpus-analysis tool*, esso consente di avere accesso ad informazioni sulla frequenza d'uso dei termini e sui diversi contesti in cui essi vengono utilizzati. A tal fine anche i corpora utilizzano la divisione di testi in *token*, termine con cui si intendono tutte le parole contenute nel corpus. A fianco a questa divisione, i corpora utilizzano anche quella in *type*, ovvero il numero di parole diverse all'interno del corpus. Per esempio, ricercando il termine "vettore", esso verrà considerato come un *type*, e il numero delle volte in cui esso appare all'interno del corpus sarà la somma di tutti i suoi *tokens*. Le *word-frequency lists*, liste di frequenza dell'uso di termini, consentono di visualizzare i *type* con accanto il numero di *token*, e di ordinare queste informazioni secondo diversi criteri, come la maggiore o minore frequenza, in ordine alfabetico o ancora secondo l'ordine di comparsa all'interno del corpus). È inoltre possibile inserire come unico *type* le forme flesse di nomi, aggettivi e verbi attraverso la lemmatizzazione di ogni *token*. Questa possibilità, se da un lato può essere utile ai fini dell'analisi testuale, dall'altro può causare fraintendimenti nel caso di parole omografe, per la cui disambiguazione servirà annotare il corpus, ovvero inserire informazioni grammaticali e sintattiche per ogni *token*.

Un altro importante strumento dei corpora, comune anche ai CAT Tools, è il *concordancer*, che consente di ricercare un termine e visualizzarlo all'interno della frase in cui esso è utilizzato. Nei sistemi di analisi dei corpora, il formato più comune di

---

<sup>7</sup> Un corpus può essere definito come: «A collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language» (Eagles, 1996, cit. in Chiari, 2008: 63) e un corpus elettronico in particolare è «a corpus which is encoded in a standardized and homogeneous way for open-ended retrieval tasks» (Eagles, 1996, cit. in Chiari, 2008: 63).

visualizzazione è definito KWIC (*Key Word In Context*), in cui tutte le occorrenze della parola ricercata sono allineate al centro dello schermo, come mostrato in figura:



voluptuous hell . Only while praying inside the sublime churches of Madrid and before the marble  
magician who whisks away the cape to reveal a sublime cinematic illusion , he only succeeds in  
Silverchair struggle to realize Johns ' sublime cinematic vision : The momentum ebbs and  
consistently dreamy quality , thanks largely to the sublime cinematography of John Seale ( Dreamcatcher  
provincial backgrounds , to become that sublime citizen . Now , post-Pulitzer , Cunningham  
Thomas Hull for a performance of Mozart 's sublime Clarinet Quintet , Another renowned chamber  
Britneyache on the Go Home site , which is a sublime clash of Public Image 's Socialist and  
of the Auvergne ) offers some of the most sublime classical music ever created . ' Bailèro  
who combined sensational virtuosity with sublime classicism ; the disciple of Balanchine  
raced into a two sets to love lead with some sublime claycourt tennis . But he began to tire  
tenuous duality - clumsy moviemaking or sublime cleverness ? - comes up in A.I. time and  
group as the architects of " Scoop " , a sublime click-infused slice of ambient techno ,  
Africa 's third largest city - offers a sublime climate , the Indian Ocean with its sandy  
with three telling shots . The first is a sublime close-up of Midori 's face shimmering through  
. Their simple but elegant faces produce sublime close-ups . But Guédiguian also gives the  
changed ever since places like the original sublime closed and the pav its just to mainstream  
of collaborative opera Listening to the sublime closing duet of Monteverdi 's L' Inconrazione

Fig. 8: Un esempio di ricerca compiuta con Concordancer in un corpus di lingua inglese<sup>8</sup>

Ordinando secondo diversi criteri i risultati della ricerca è possibile ottenere informazioni diverse sull'utilizzo di espressioni e parole. Per esempio, ordinando i risultati alfabeticamente in base alla parola precedente o successiva a quella ricercata, è possibile rilevare eventuali espressioni ricorrenti, termini composti, ecc. Le ricerche condotte in un corpus parallelo produrranno invece frasi e periodi contenenti la parola ricercata, affiancati dalla frase o dal periodo corrispondente nella seconda lingua. In questo modo si potranno confrontare i due testi e verificare la corrispondenza di determinate espressioni nell'altra lingua. I corpora bilingui sono solitamente bidirezionali, quindi si può svolgere una ricerca in ognuna delle due o più lingue.

Un ultimo e importante strumento dei corpora è rappresentato dalla ricerca delle collocazioni (Seretan, 2009 e 2010), ovvero parole che vengono utilizzate sovente insieme, formando dei cluster. Può essere il caso di espressioni idiomatiche, termini composti, costrutti sintattici. Le collocazioni vengono distinte dai casi di parole che si trovano vicine per motivi funzionali o dettati dal caso tramite un generatore di collocazioni che utilizza una formula di mutua informazione (*mutual information formula*). Tramite essa è possibile calcolare il rapporto tra la probabilità di incontrare due

<sup>8</sup> Tratto dal Collins Cobuild Corpus (stralcio disponibile sul web)

parole insieme e la probabilità di incontrarle individualmente. Maggiore la frequenza con cui le parole appaiono insieme all'interno del corpus, maggiore sarà la possibilità che esse formino una collocazione. Per non considerare collocazioni anche i composti articolonome o nome-preposizione occorrerà aggiungere delle annotazioni fraseologiche e sintattiche, ovvero informazioni che aiutino a distinguere le varie parti del discorso (*ibid.*).

#### **1.4.5. Lo *specialized discourse***

Infine, per comprendere appieno l'utilizzo dei CAT Tools un ulteriore importante aspetto da considerare è la tipologia testuale alla quale essi rispondono meglio, ovvero i linguaggi specialistici e tecnici, che vengono studiati sotto il nome di *specialized discourse*. I linguaggi specialistici sono molteplici e ognuno di essi ha delle caratteristiche particolari e delle specifiche tipologie testuali afferenti ad esso. Gli elementi che tuttavia li accomunano sono un alto tasso di standardizzazione, una funzione comunicativa ben definita, un contesto di utilizzo ed un target di fruitori delineato (la *discourse community*), l'utilizzo di terminologia propria e di forme stilistiche specifiche, un registro spesso formale e impersonale (Gotti, 2008). I criteri dominanti che regolano questi linguaggi sono quello economico (esprimere i concetti in maniera concisa), quello della precisione (esprimere i concetti in dettaglio e senza trascurare alcuna informazione) e dell'appropriatezza (utilizzare la giusta tipologia testuale e il registro più consono alle finalità del testo)<sup>99</sup>. Questi tre criteri dovranno essere rispettati al fine di ottenere il massimo dell'efficacia nella comunicazione. Un traduttore che debba tradurre testi appartenenti ad un particolare ambito settoriale deve conoscere molto bene le caratteristiche del linguaggio specialistico corrispondente sia nella lingua di partenza che di arrivo, e padroneggiarne lo stile per produrre una traduzione che sia perfettamente rispondente alle aspettative dei destinatari e alle caratteristiche del settore dal quale il testo verrà recepito. Per esempio, un traduttore che intenda lavorare in ambito medico non dovrà solamente procurarsi adeguati glossari e dizionari settoriali, ma dovrà anche conoscere le diverse tipologie testuali tipicamente prodotte in ambito medico, e lo stile, le espressioni comuni, le prassi linguistiche e sintattiche che caratterizzano questo particolare ambito del linguaggio. Bisognerà tenere in mente la finalità del testo da

---

<sup>99</sup> Sager *et al*, 1980, cit. in Gotti, 2008: 30

tradurre e il pubblico a cui esso è destinato. Il grado di formalità e di specializzazione sarà più alto se l'utilizzo del testo sarà *ad intra*, ovvero destinato alla comunità dei medici ed esperti del settore, mentre sarà minore nel caso di testi esplicativi destinati ad un pubblico più ampio e meno competente nel settore. Una così grande precisione dal punto di vista stilistico e terminologico non può che beneficiare dell'utilizzo di memorie e glossari da consultare ed utilizzare nel corso del lavoro, che garantiscano il rispetto delle regole e dei parametri che regolano il dato linguaggio che si sta traducendo. La rigidità delle regole che ordinano i linguaggi specialistici, quindi una certa prevedibilità in fatto di costrutti e stile, e l'elevato grado di ripetitività interna rispetto ad altre tipologie e generi testuali come i testi letterari o creativi, sono le caratteristiche che rendono questi linguaggi particolarmente adatti a sfruttare al meglio le risorse offerte dai CAT Tools.

È tipica dei linguaggi specialistici la presenza di espressioni standardizzate e vere e proprie formule, come nel caso dei documenti medici, dei termini e condizioni che regolano la maggior parte dei contratti di compravendita, dei manuali di uso e manutenzione. La ripetitività interna e la ricorsività delle espressioni rende utile il ricorso alle memorie di traduzione per il duplice beneficio di riutilizzare porzioni di testo già tradotto (beneficio economico) e di far sì che venga garantita la coerenza terminologica all'interno del documento (beneficio in termini di qualità). I traduttori professionisti utilizzano TM e glossari specifici per i diversi settori nei quali lavorano maggiormente, e alle volte costruiscono materiali di riferimento ancora più settoriali per ambiti che, pur ricadendo all'interno dello stesso campo hanno una terminologia o una fraseologia molto tecnica e precisa da richiedere delle risorse più che affidabili e dedicate. La capacità di saper offrire alla clientela questo valore aggiunto in termini di affidabilità e di specializzazione fa la differenza all'interno di un mercato altamente competitivo ed esigente.

## Capitolo 2: Tipologie di strumenti

### 2.1 *Translation Memories*

Dopo l'ampia introduzione del capitolo precedente, si procederà adesso a descrivere in dettaglio le varie tipologie di strumenti e le loro funzioni. Per cercare di rendere il quadro sufficientemente ampio si farà riferimento a diversi CAT Tools, nello specifico a SDL Trados nelle sue versioni del 2007 e del 2014<sup>10</sup> e a MemoQ<sup>11</sup> (in 2.2), che rappresentano le scelte principali al giorno d'oggi tra l'ampia offerta di suite e programmi di *computer-aided translation*. Altri CAT Tools tra i più utilizzati sono SDLX, Wordfast, Atril Déja Vu, Across e tra gli open-source spicca OmegaT.

#### 2.1.1. Le TM

I CAT Tools si presentano spesso come sistemi integrati che includono diversi programmi e funzionalità. Il “Workbench” è a tutti gli effetti il nuovo banco di lavoro del traduttore, che gli permette di trovare a portata di pochi clic tutto ciò di cui ha bisogno. I Workbench comprendono quasi sempre memorie di traduzione, strumenti di allineamento, strumenti di pre-analisi dei testi, applicativi per la creazione di glossari e sistemi di gestione terminologica, nonché una serie di comandi per la revisione e il controllo della traduzione completa. Il cuore dei CAT Tools, tuttavia, è rappresentato dalle TM (*Translation Memory*), le memorie di traduzione. Diffuse a partire dalla metà degli anni '90 esse sono state definite come degli archivi di testi multilingui segmentati, allineati e classificati che consentono la memorizzazione ed il recupero di segmenti testuali multilingui allineati a partire da diversi criteri di ricerca (EAGLES 1996, cit. in Quah 2006). Diversamente dai sistemi di traduzione automatica, le memorie di traduzione non propongono soluzioni prodotte dall'intelligenza artificiale della macchina, ma lasciano al traduttore il compito di tradurre e “si limitano” a suggerire porzioni di testo

---

<sup>10</sup> Le informazioni sono tratte dai manuali disponibili sul sito [www.sdl.com](http://www.sdl.com)

<sup>11</sup> Il manuale di MemoQ è disponibile all'indirizzo [https://memoq.com/Kilgray/media/Download/Guides/memoQ\\_QuickStartGuide\\_6\\_5\\_EN.pdf](https://memoq.com/Kilgray/media/Download/Guides/memoQ_QuickStartGuide_6_5_EN.pdf)

precedentemente tradotte in base al loro grado di similarità con il nuovo testo da tradurre, e a salvare ogni nuovo segmento che viene tradotto, allineando automaticamente testo di partenza e testo di arrivo. In questo modo il traduttore può creare la propria memoria di traduzione man mano che avanza nel proprio lavoro.

I programmi di creazione delle memorie di traduzione consentono ogni possibile combinazione linguistica poiché tali strumenti non si basano su un'analisi di tipo linguistico ma sulla divisione in frasi e paragrafi e su degli algoritmi di calcolo del grado di somiglianza che tengono conto del numero di parole uguali all'interno delle frasi. Essi sono quindi *language-independent* e possono inoltre supportare diversi alfabeti e script. La memoria può essere virtualmente espandibile all'infinito, e a seconda delle necessità dell'utente può essere più o meno settorializzata. Un traduttore può infatti decidere di lavorare con la stessa TM per tutti i clienti e per tutti gli ambiti linguistici, o può piuttosto creare ed utilizzare TM diverse in base ai clienti e ai progetti ai quali lavora. Quest'ultima possibilità è solitamente preferita dai professionisti, poiché essa consente una maggiore garanzia di coerenza rispetto al contesto in cui la traduzione si situa, l'utilizzo della corretta terminologia specialmente laddove essa sia altamente specialistica e puntuale, il mantenimento degli stessi accorgimenti stilistici adottati per un dato settore. La consultazione delle TM può avvenire in automatico o su stimolo del traduttore nel corso del lavoro, e le ricerche che tramite essa si possono compiere riguardano non solo intere frasi o segmenti, ma anche singoli termini o espressioni, poiché la memoria si interfaccia anche con gli strumenti di gestione terminologica.

Traducendo con un CAT Tool, la procedura prevede la selezione di una TM preesistente o la creazione di una nuova, che all'inizio del lavoro sarà quindi vuota e da popolare. In alcuni casi, quando si lavora in team sotto la supervisione di un project manager, questi fornirà a tutti i traduttori impegnati nel progetto una TM da utilizzare, contenente già in parte la terminologia e le espressioni chiave, in molti casi concordate con il committente.

### **2.1.2. Fuzzy match, Exact match**

Il CAT Tool dividerà il testo in porzioni, solitamente corrispondenti a frasi e periodi, o in base alla punteggiatura. Tali porzioni o segmenti di testo verranno confrontate con le doppiette di testo in lingua *source* e *target* presenti in memoria, dette *Translation Units*, per ricercare eventuali corrispondenze e soluzioni utilizzabili all'interno della nuova traduzione.

La consultazione di una TM può produrre due tipi di risultati, un *perfect matching* o un *fuzzy matching*<sup>12</sup>. Il primo, anche noto come *exact matching*, viene prodotto quando un nuovo segmento da tradurre è identico al 100% ad un segmento presente in memoria, come da immagine:

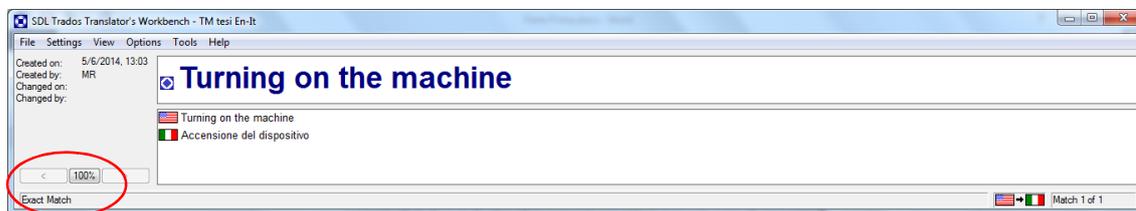


Fig. 9: Un esempio di exact matching in Trados Workbench 2007

La schermata sopra riportata rappresenta il modo in cui Trados 2007 mostra all'utente i risultati che ha trovato in memoria. In basso a sinistra viene segnata la percentuale di similarità, in questo caso 100%, mentre nelle due finestre principali viene visualizzata in alto la frase da tradurre, ed in basso la *Translation Unit* recuperata dalla memoria. Malgrado la somiglianza pari al 100%, questa versione del programma non inserirà la frase automaticamente, ma aspetterà che sia l'utente a confermare o meno l'utilizzo della soluzione già presente in memoria. Le versioni più recenti di Trados, invece, offrono l'opzione di inserimento diretto degli *exact matching*, come si vedrà più avanti. Siamo invece in presenza di un *fuzzy matching* quando un nuovo segmento da tradurre ha un elevato grado di somiglianza con uno o più segmenti salvati in memoria. Può bastare anche solo una differenza minima come la punteggiatura, o una somiglianza al 99%, affinché un risultato venga segnato come *fuzzy*.

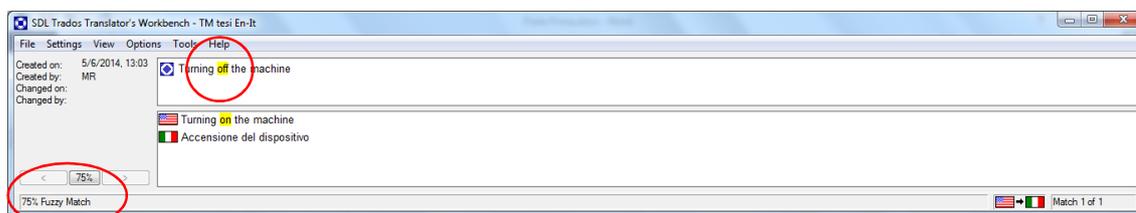


Fig. 10: Un esempio di fuzzy matching in Trados Workbench 2007

<sup>12</sup> La differenza tra i due tipi di match è la stessa che c'è tra le espressioni *translation-equivalence pair* e *translation-equivalence candidate* (Tufis et al, 2004:164), in relazione all'estrazione automatica di equivalenti traduttivi da un corpus parallelo (con il quale, si vedrà, la TM ha molto in comune).

Come nel caso precedente, il programma proporrà la soluzione già immagazzinata, evidenzierà in giallo l'elemento che differisce tra le due frasi (nell'esempio sopra riportato "Turning off" invece di "Turning on"), segnerà la percentuale di somiglianza e, nel caso di più soluzioni che potrebbero tornare utili all'utente (per esempio se in memoria si fosse trovata una frase come "Turning off the printer" invece che "the machine"), queste potranno essere visualizzate cliccando sulle freccette ai lati dell'indicatore della percentuale di matching. L'utente ha la possibilità di fissare a propria discrezione la percentuale di matching a partire dalla quale desidera visualizzare i suggerimenti estratti dalla memoria. Sebbene abbassando la soglia minima di matching sarà più facile ottenere dei riscontri, in generale la buona pratica è quella di non fissare un limite troppo basso, poiché i risultati ottenuti potrebbero essere solo lontanamente simili e non contribuire in maniera utile a trovare una soluzione traduttiva. La percentuale impostata di default è generalmente 65 – 70%, valori più alti avrebbero dal canto loro lo svantaggio di fornire pochissimi risultati all'utente. La funzione di fuzzy matching consente al traduttore di valorizzare al massimo il materiale già tradotto, con un notevole risparmio in termini di fatica e di tempo, ma anche a vantaggio della coerenza di espressione e nell'uso della terminologia in documenti appartenenti allo stesso ambito o destinati allo stesso cliente. Laddove la ricerca non produca dei match, è comunque possibile compiere la ricerca di porzioni più piccole di testo all'interno della memoria, tramite lo strumento denominato *Concordance Search*, di cui viene mostrato un esempio nella figura seguente. La ricerca produrrà come risultato la lista dei segmenti di testo contenenti l'espressione o il frammento di testo selezionato, che sarà evidenziato in giallo. In tal modo se ne potrà osservare l'utilizzo all'interno dei diversi contesti linguistici e decidere se la traduzione adottata sia adatta al caso. Il *Concordance Search* fornisce quello che Bowker e Fisher (2010: 62) definiscono *Sub-segment match*.

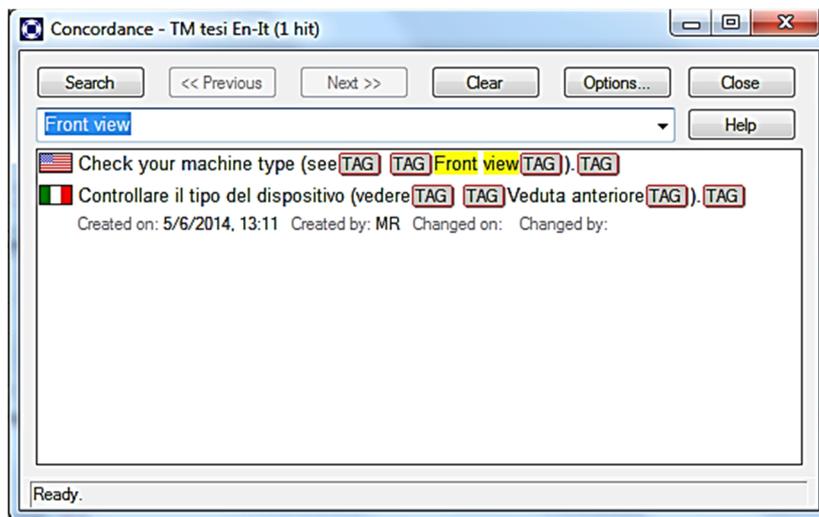


Fig. 11: Un esempio di Concordance search dell'espressione "Front view", in Trados 2007

Altre utili funzioni per il riutilizzo di materiale già tradotto sono il *Translate to fuzzy* e l'*Autopropagate*, entrambe non automatiche ma avviate dal software solo se richiesto dall'utente. Il primo scorre tutto il documento inserendo automaticamente nel testo tutti gli exact match che incontra, fermandosi laddove trovi un fuzzy match o una frase interamente da tradurre. Il secondo, disponibile in versioni successive di Trados, oltre ad inserire automaticamente gli exact match presenti in memoria, ogni volta che un nuovo segmento viene tradotto si accerta della presenza di eventuali frasi uguali al 100% ed estende la traduzione appena compiuta anche a tali segmenti. In caso di modifiche al segmento già tradotto, inoltre, la modifica verrà applicata a tutti i segmenti uguali. Prima di iniziare una nuova traduzione, utilizzando la TM è possibile analizzare il testo per verificare in termini percentuali la quantità di testo già presente in memoria, la quantità di fuzzy match e infine la quantità di testo interamente da tradurre. Questa funzione, nota come *Analyze*, consente al traduttore di comprendere in anticipo il carico di lavoro che lo attende e di fornire al cliente un preventivo preciso o comprendere quanto verrà retribuito dall'agenzia con cui lavora, data la tendenza ormai affermata sul mercato di offrire tassi di pagamento diversi per i no match, i fuzzy e gli exact matching. La funzione di *Analyze* può essere applicata sia a documenti singoli che in batch.

### 2.1.3. L'allineamento

Per popolare una TM è possibile come si è visto memorizzare i segmenti man mano che si traduce, opzione che renderà la stessa TM utile solo dopo aver tradotto un

considerevole numero di testi, o importare memorie condivise dai project manager tramite la funzione *Import*. Al fine di rendere le TM facilmente scambiabili, è stato sviluppato un formato di interscambio comune tra tutti i CAT Tools, il *.tmx*, che viene poi trasformato dal Workbench nel formato proprio del CAT Tool in uso. Anche l'opzione inversa è possibile, ovvero l'esportazione della TM dal Workbench nel formato di interscambio. È buona prassi infatti restituire al Project Manager la memoria con la quale si è lavorato, aggiornata con i nuovi segmenti prodotti. Infine, una TM può essere popolata anche a partire da testi già interamente tradotti o disponibili nelle due lingue, tramite la funzione di allineamento già richiamata nel parlare dei corpora. Tale funzione, che nei diversi programmi prende il nome di *Alignment* o *Align*, analizza i testi e ne identifica la divisione in frasi e periodi. Successivamente attua un collegamento frase per frase sulla base del riconoscimento della stessa punteggiatura o livello di testo (punto elenco, sottopunto ecc.) ma anche sulla base della presenza di terminologia nota nel caso in cui il programma si interfacci con gli strumenti di gestione terminologica, o grazie alla presenza di numeri, date ecc. al fine di garantire la maggiore correttezza possibile di allineamento.

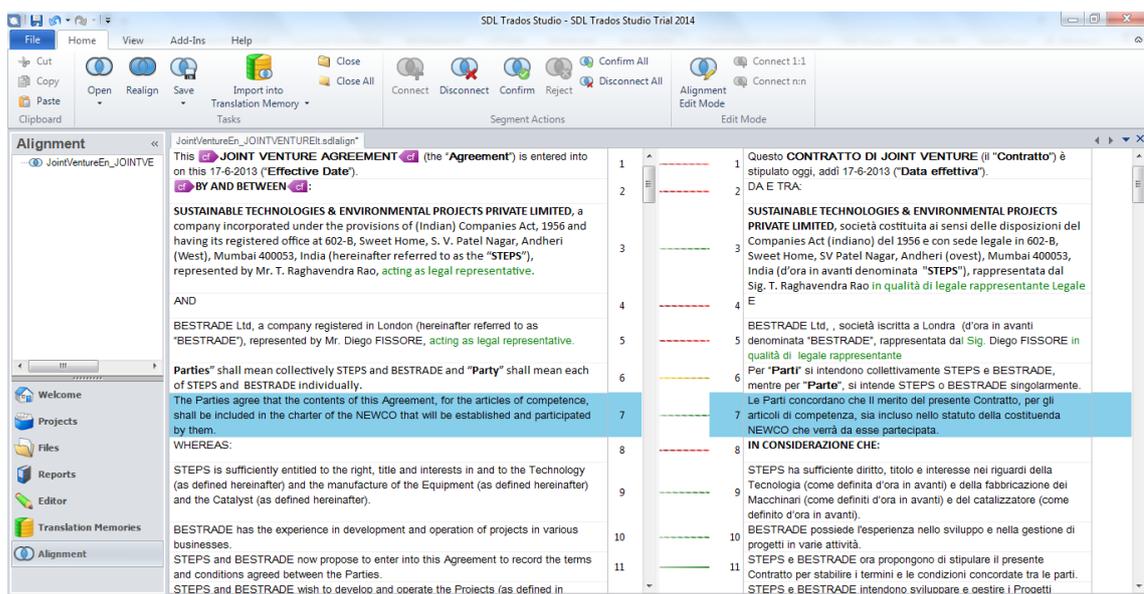


Fig. 12: La schermata dello strumento di Allineamento di Trados Studio 2014

Il traduttore potrà correggere manualmente eventuali errori di allineamento. Una volta terminato il lavoro e verificato che tutti i segmenti di testo siano stati allineati correttamente con i corrispettivi nella seconda lingua, l'utente potrà confermare e trasformare il risultato di tale processo in *Translation Units* da inserire nella propria TM

ed utilizzare per traduzioni successive. Questo processo può tornare utile ai traduttori all'inizio della loro carriera, che possono beneficiare delle TM condivise dai project manager per cominciare a creare il proprio bagaglio di risorse, o sfruttare i materiali bilingui attendibili disponibili su internet in siti specializzati nei vari ambiti in cui essi intendono lavorare. Anche i traduttori con esperienza e che iniziano dopo tempo ad utilizzare un CAT Tool possono avvalersi di questa funzione; con l'allineamento potranno salvare in memoria le traduzioni compiute senza l'ausilio del software, e dare quindi un valore aggiunto al lavoro precedentemente svolto.

#### **2.1.4. Sistemi alternativi**

Prima di passare avanti, è interessante fare riferimento anche a un paio di sistemi non altrettanto noti quanto Trados, e che utilizzano un approccio diverso alla divisione in segmenti per la creazione delle TM. I principali CAT Tools infatti svolgono l'operazione di segmentazione e allineamento dei testi a monte, per poi salvare ogni singolo segmento come TU. Invece, alcuni sistemi quali MultiTrans (<http://multicorpora.com/>) e LogiTrans (<http://www.terminotix.com/>) adottano un approccio diverso salvando in memoria l'intero testo insieme al suo corrispettivo nella lingua di arrivo, e solo dopo averlo salvato ne riconoscono la scomposizione in paragrafi e compiono l'allineamento. I vantaggi principali di tale approccio sono una più veloce creazione di grandi TM e la possibilità di visualizzare il testo nella sua interezza, senza la divisione in segmenti che può rendere la fruizione del testo stesso troppo frammentaria e disarticolata. Ancora più particolare è l'approccio adottato da STAR Transit (<http://www.star-group.net/ITA/transit-nxt/transit.html>), che non si avvale di una vera e propria TM per l'immagazzinamento e il recupero dei testi tradotti, ma trasforma il PC dell'utente in una grande *Translation Memory* virtuale, dalla quale il traduttore selezionerà i documenti rilevanti ai fini della propria traduzione, che verranno convertiti dal software in formato *.xml* al fine di poterli confrontare con il testo da tradurre, alla ricerca di risultati utili.

Anche il metodo di ricerca delle informazioni all'interno delle TM può essere diverso. Nella maggior parte dei casi esso viene compiuto tramite riconoscimento delle stringhe di caratteri, mentre alcuni software come SIMILIS o Masterin affiancano a questa tecnica di matching anche delle informazioni linguistiche che consentono di compiere un'analisi per rinvenire anche le unità sintattiche (chunk) alle quali vengono allegate delle annotazioni grammaticali (Lagoudaki 2006: 4). Ciò rende possibile, nel caso in cui non

si trovi alcun match per un segmento, che il software crei un fuzzy match andando a ricercare termini e strutture sintattiche in segmenti diversi. Se da un lato questa opzione consente di avere il vantaggio di offrire informazioni grammaticali utili a disambiguare o comprendere meglio i segmenti da tradurre, dall'altro essa rende il software *language-dependent*, ovvero capace di creare TM solo per quelle lingue per le quali siano presenti adeguate informazioni linguistiche e grammaticali. Questo fa venir meno la possibilità di applicare il sistema a qualsiasi combinazione linguistica.

## 2.2 *L'ambiente di lavoro*

### 2.2.1. *La Workstation*

I primi modelli di CAT Tool, come ad esempio Wordfast o lo stesso Trados, funzionavano come Macro da attivare all'interno del programma di videoscrittura, quindi consentivano di modificare il testo direttamente nel suo formato originario e di visualizzare contemporaneamente il testo originario e di seguito la sua traduzione via via che veniva svolta. Tale modalità di visualizzazione viene ancora offerta come possibile applicazione da diversi CAT Tools, tuttavia essa può risultare per diversi aspetti alquanto scomoda, intanto per il rischio di compromettere involontariamente il testo originario, ed in secondo luogo per i casi in cui vi sia una particolare formattazione testuale che non consenta un'agevole visualizzazione del testo. Inoltre rende possibile l'applicazione del software solamente a pochi formati di file, principalmente i *.doc*. Si andò via via preferendo quindi l'adozione di un ambiente di lavoro dedicato, all'interno del quale importare il file da tradurre convertendolo in un formato adatto ad essere tradotto dal CAT Tool mantenendo tuttavia inalterata la formattazione o il layout, e visualizzando in maniera comoda e funzionale sia il testo originario che quello man mano prodotto dal traduttore. Questa modalità consentiva una più agevole elaborazione dei testi per numerosi formati file, e al giorno d'oggi è possibile importare anche formati molto specifici come i PDF non protetti, Interleaf, InDesign, QXP, PageMaker ecc. L'immagine di seguito riportata mostra l'ambiente di lavoro di Trados 2007.

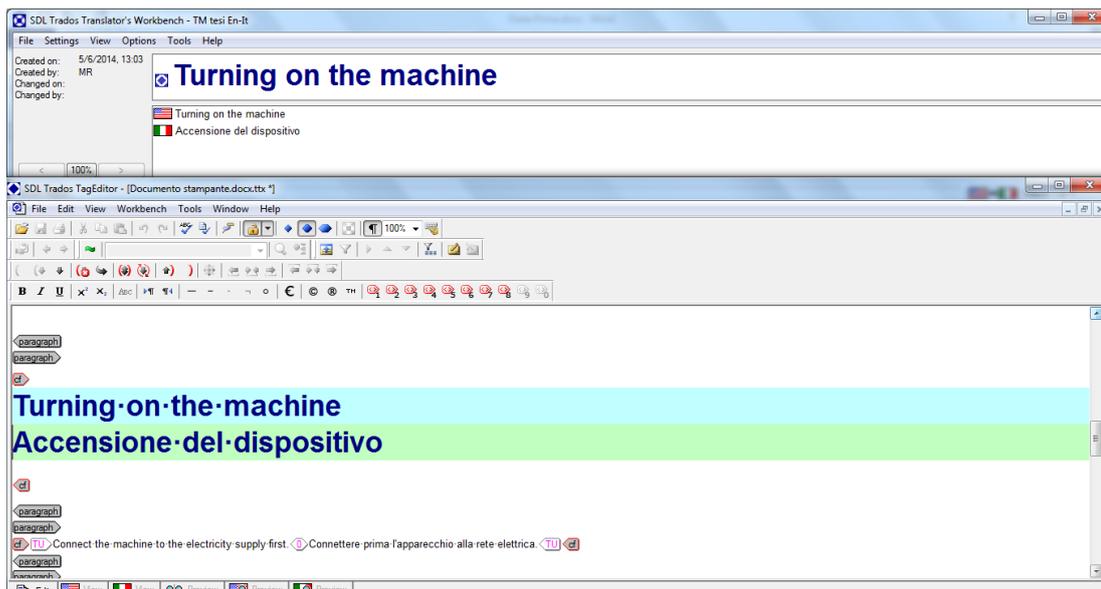


Fig. 13: Il TagEditor di Trados 2007

La parte superiore è il Workbench, già mostrato precedentemente, ovvero la finestra che mostra i risultati ottenuti interrogando la memoria, con la specifica della percentuale di matching. La parte inferiore, denominata *TagEditor*, è la finestra all'interno della quale il traduttore opera sul testo. È possibile distinguere il testo segmentato in frasi, e degli elementi grigi a forma di freccia denominati *tag* o *placeables*, che forniscono informazioni relative alla formattazione e al layout del testo, quindi la divisione in paragrafi, l'utilizzo di corsivo, grassetto, testo di diversa grandezza, carattere o colore, la presenza di immagini o collegamenti ipertestuali ecc. L'utilizzo di questi *placeables* solleva il traduttore dalla necessità di modificare il formato o di ricreare l'impaginazione originaria del testo. Tutte queste informazioni infatti verranno trasposte all'interno del testo tradotto grazie ai *placeables*, alcuni dei quali vengono inseriti automaticamente, mentre per altri (quelli che si trovano all'interno dei segmenti) sarà l'utente a dovere aver cura di riportarli in maniera esatta e nel giusto ordine. Nell'immagine si distingue una frase evidenziata in azzurro, ovvero il segmento corrente sul quale si sta lavorando, con la traduzione sottostante evidenziata invece in verde. Questo è il colore col quale vengono contrassegnati gli exact matching quando vengono rilevati dal programma nel momento in cui si seleziona il segmento da tradurre. I fuzzy matching vengono invece contrassegnati in giallo chiaro.

L'immagine seguente riporta l'ambiente di lavoro di Trados 2014.

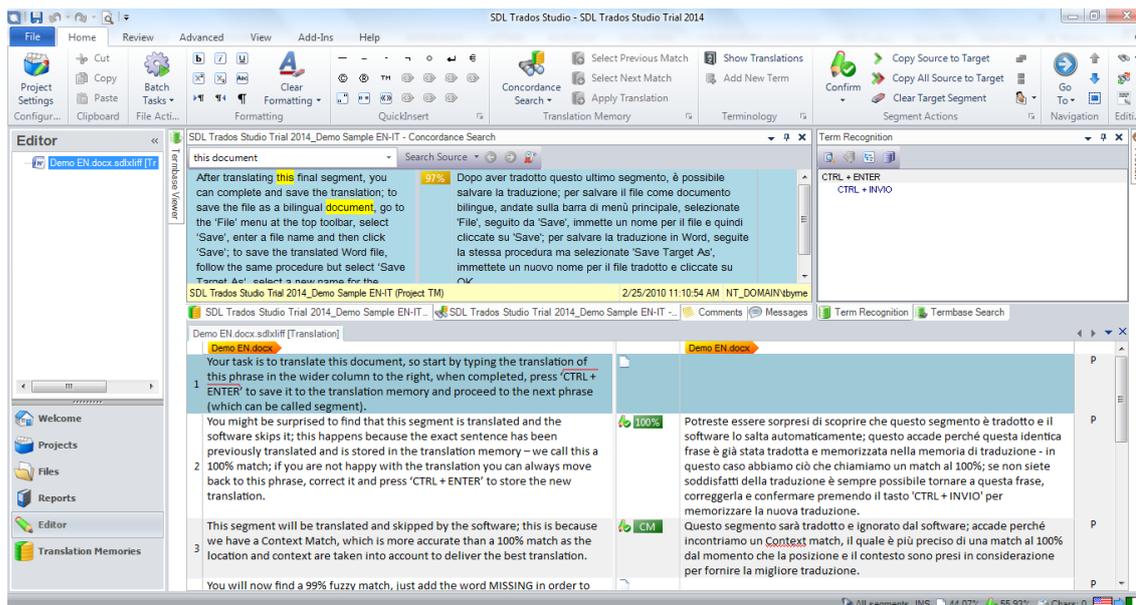


Fig. 14: L'ambiente di lavoro in Trados Studio 2014

Il testo originario, sempre diviso in segmenti, è in questo caso affiancato al testo tradotto. Questa nuova visualizzazione si è dimostrata più funzionale e comoda perché consente di avere una panoramica meno frammentata dei due testi, cosa che aiuta il traduttore a orientarsi meglio all'interno del documento e a percepire il testo da tradurre come un tutt'uno, a favore di una visione d'insieme che veniva a mancare a causa della sensazione di frammentarietà comportata dalla divisione in segmenti (Mossop, 2006 e Lagoudaki, 2006). La memoria di traduzione trova spazio al di sopra del testo, mentre la divisione tra exact e fuzzy matching è specificata nella colonna che divide i due testi. Nell'immagine, la frase è contrassegnata non da una percentuale, ma da un'abbreviazione, CM, che sta per *Context Match*. Con essa si intende un exact match che si trova contestualmente accanto ad altri match. La presenza di più frasi ravvicinate delle quali è presente un match rappresenta una maggiore garanzia di correttezza e conformità della traduzione, che il traduttore potrà adottare con un alto grado di sicurezza. Proprio per questo motivo esse vengono inserite automaticamente dal software e direttamente confermate. La nuova visualizzazione di Trados Studio 2014, a dire il vero già adottata a partire dalla versione del 2009, ha anche di buono che tutte le risorse trovano spazio all'interno di un unico ambiente, rendendo più agevole la visualizzazione e l'armonizzazione delle varie parti. Tutti i riquadri nei quali trovano spazio tali risorse sono inoltre regolabili e personalizzabili dall'utente, a seconda della visualizzazione che gli è più congeniale. Il

menù a sinistra consente inoltre di avere un accesso immediato alle funzioni di project management come quelle per la gestione dei glossari e la visualizzazione dei report e del grado di avanzamento dei documenti del progetto, segnata in percentuale alla sezione “Files”.

La scelta dell’ambiente unico di traduzione è stata anche preferita dagli sviluppatori di MemoQ come si vede nell’immagine sotto riportata. Relativamente alla visualizzazione del testo da tradurre, l’utente potrà decidere al momento dell’installazione se adottare quella affiancata o quella con le frasi source e target uno di seguito all’altro. Si potranno inoltre selezionare le risorse che si vorranno visualizzare nei riquadri a destra, tra i quali glossari online come l’Euro Termbank o servizi di traduzione automatica offerti da diversi siti.

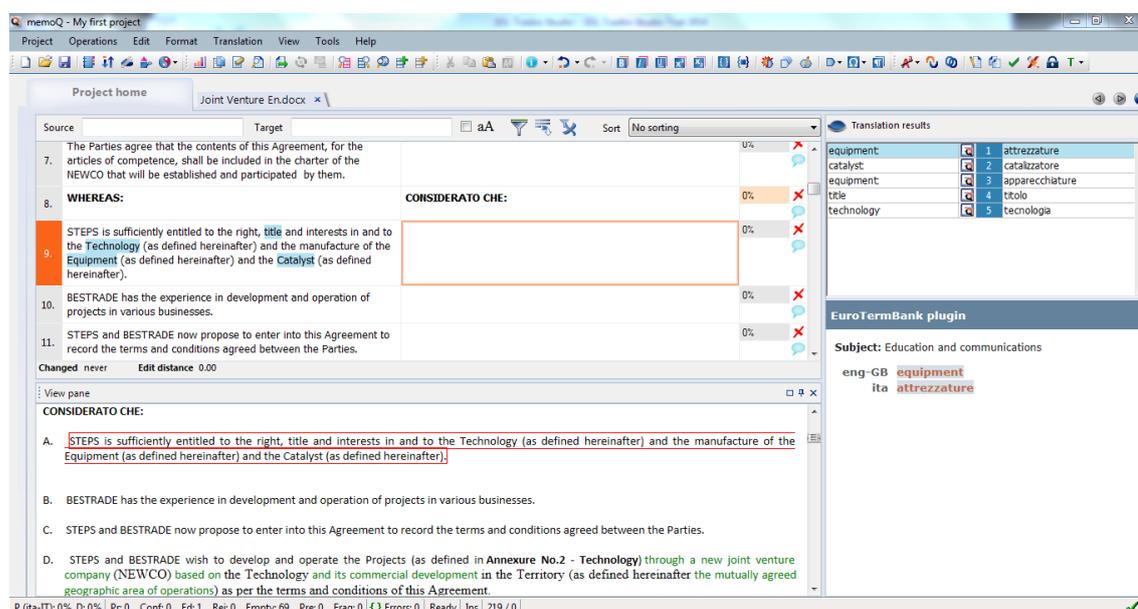


Fig. 15: L’ambiente di lavoro in MemoQ

### 2.2.2. Le funzioni disponibili

Il TagEditor, così come gli ambienti di lavoro degli altri CAT Tools, offre numerose funzioni di ricerca, analisi e controllo, per monitorare sotto tutti gli aspetti l’evoluzione del lavoro in ogni momento. È da qui infatti che vengono avviate le interrogazioni della TM, sia la ricerca dei match che la “Concordance search”, nonché la consultazione dei glossari selezionati. Esso offre inoltre numerose utili opzioni di lavoro. Procedendo nella traduzione, per esempio, è possibile decidere di salvare direttamente in

memoria la nuova TU creata o di lasciarla in stand-by per ritornare in un secondo momento e rivederla in caso di dubbi sulla correttezza delle scelte compiute. Una volta confermata, essa verrà immagazzinata nella memoria di traduzione. Viene offerta anche la possibilità di aggiungere commenti nel caso in cui il traduttore voglia sottoporre alcuni segmenti in particolare all'attenzione del revisore, o al contrario per consentire a quest'ultimo o al proofreader di segnalare eventuali errori, sviste o imprecisioni a chi ha tradotto. Sono inoltre offerte diverse possibilità di visualizzazione, utili a controllare il progresso della traduzione e il corretto utilizzo dei *placeable* relativi alla formattazione. Si può visualizzare infatti, direttamente nell'ambiente di lavoro o in una finestra a parte, solo il testo originario o solo quello tradotto, o entrambi contemporaneamente in due finestre sovrapposte. Questa risorsa ha concretamente la stessa funzione dell'anteprima di stampa in Word. Molto utile può anche essere la possibilità di espandere e restringere i segmenti. Le impostazioni di default dividono il testo in segmenti sulla base della punteggiatura o della divisione in paragrafi. Tuttavia in alcuni casi il traduttore può voler restringere ulteriormente il segmento o ampliarlo in base, per esempio, alla difficoltà di far combaciare la traduzione col testo originario, o per salvare in memoria porzioni più piccole di testo che possano quindi essere più facilmente riutilizzabili. Avere segmenti più piccoli, inoltre, può agevolare il ritrovamento di match, mentre segmenti troppo grandi lo rendono più difficile.

Nel caso di testi in cui vengano utilizzati spesso termini composti lunghi come il nome di società o prodotti, o nomenclatura scientifica estesa, per risparmiare tempo ed essere certi che tali termini ricorsivi siano sempre utilizzati correttamente, oltre all'utilizzo dell'inserimento da glossario è anche possibile usufruire dello strumento "Auto Text", che consente di creare una lista di parole o espressioni che verranno inserite nel testo semplicemente digitando le prime lettere e premendo invio. Il traduttore potrà aggiungere o eliminare espressioni di "testo automatico" in base alle esigenze del documento sul quale sta lavorando. Infine, tutti i CAT Tools offrono un'ampia gamma di strumenti di revisione e quality check, da utilizzare durante la traduzione o alla fine di essa. I parametri che è possibile monitorare e verificare sono molteplici, dal semplice *spell checker* al controllo del corretto utilizzo dei tag, della punteggiatura e delle maiuscole, dei numeri contenuti all'interno del testo, di eventuali ripetizioni nei segmenti target e di segmenti lasciati per sbaglio non tradotti o tradotti solo in maniera parziale. È possibile inserire

liste di parole corrette e sbagliate in modo da riscontrare l'utilizzo di alcune espressioni errate e sostituirle facilmente con la corrispondente espressione corretta, e infine è possibile verificare il corretto utilizzo della terminologia attraverso il confronto automatico con i glossari di riferimento. Esistono inoltre degli strumenti specifici per alcuni formati di testo come Excel e per i casi in cui si lavori in team. Questo pacchetto di opzioni di verifica, in TagEditor selezionabili tramite l'opzione "Plug-ins", rappresenta un valore aggiunto e un significativo aiuto per il traduttore nel garantire un lavoro di elevata qualità e professionalità.

Un'ultima annotazione di grande rilevanza riguarda come sopra accennato l'introduzione in diversi CAT Tools come SDL Trados Studio e MemoQ, della possibilità di avvalersi della traduzione automatica per compiere laddove ritenuto necessario una pre-traduzione del testo, in modo da accelerare il lavoro avendo una bozza di traduzione prodotta dalla macchina, che il traduttore dovrà rivedere e correggere. Molte agenzie ricorrono a questa possibilità per una questione di risparmio sui prezzi, dato che i costi di revisione sono inferiori a quelli per la traduzione, per questo diversi produttori di software hanno optato per questa forma di ibridazione tra traduzione automatica e assistita. In realtà, il ricorso alla traduzione automatica è visto di cattivo occhio dai traduttori professionisti (cf. 7.1), poiché la scarsa qualità dei testi da essa prodotti richiede comunque un lavoro molto impegnativo da parte del traduttore, molto più di una reale revisione svolta su documenti tradotti in maniera tradizionale. I professionisti vedono in tale possibilità più uno stratagemma delle agenzie per sottopagare il loro lavoro che un effettivo aiuto o un beneficio in termini di semplificazione e accelerazione del processo traduttivo.

### 2.3 *Strumenti di gestione terminologica*

Per i traduttori professionisti specializzati in materie e settori ad elevata vocazione tecnica, la terminologia è una componente cruciale del loro lavoro. Secondo Arntz e Pitch (cit. in Austermühl 2001), i traduttori impiegano circa il 75% del loro tempo di lavoro in ricerche terminologiche ma, d'altro canto, i tempi ristretti del mercato non sempre consentono la spesso lunga e approfondita ricerca del termine corretto; per questo è fondamentale che i traduttori possano contare su supporti efficaci che permettano di risparmiare tempo ma anche di garantire un corretto utilizzo della terminologia.

Numerose sono le risorse disponibili in tal senso, dai dizionari cartacei generici o specialistici, alle loro versioni su CD, fino ai dizionari e glossari disponibili gratuitamente online. Tuttavia, tali risorse, sebbene utili, sono esclusivamente consultabili e non danno al traduttore la possibilità di personalizzare i risultati per future ricerche, o di creare glossari specifici per clienti o progetti, o condividere in team i risultati in maniera facile e funzionale. Gli strumenti di gestione terminologica consentono di compiere queste e altre utili funzioni, sono quindi essenziali per un accurato uso e mantenimento della terminologia ed è fondamentale che il traduttore ne padroneggi l'uso per sfruttarne al meglio le potenzialità, e inizi molto presto nell'arco della sua carriera a costruire i propri glossari che, insieme alle memorie di traduzione, rappresenteranno il frutto più prezioso del suo lavoro.

### **2.3.1. I *terminology management tools***

La funzione principale degli strumenti di gestione terminologica, o *terminology management tools*, è quella di creare e mantenere un glossario, riutilizzarne la terminologia, disambiguare tra equivalenti multipli, fornire all'utente informazioni contestuali e rendere agevole lo scambio di terminologia tra più traduttori (Bowker e Fisher, 2010). All'inizio di un nuovo progetto, è prassi comune che il cliente o il project manager fornisca ai traduttori un elenco dei termini principali da utilizzare e ai quali attenersi scrupolosamente. Questo incipiente glossario viene poi ampliato e condiviso con il contributo di tutti in modo che vengano compiute le stesse scelte terminologiche all'interno del progetto. La figura sotto riportata è un esempio di discussione tra traduttori e project manager (o terminology manager, se la figura specifica è prevista all'interno del team), relativamente alle scelte terminologiche da compiere:

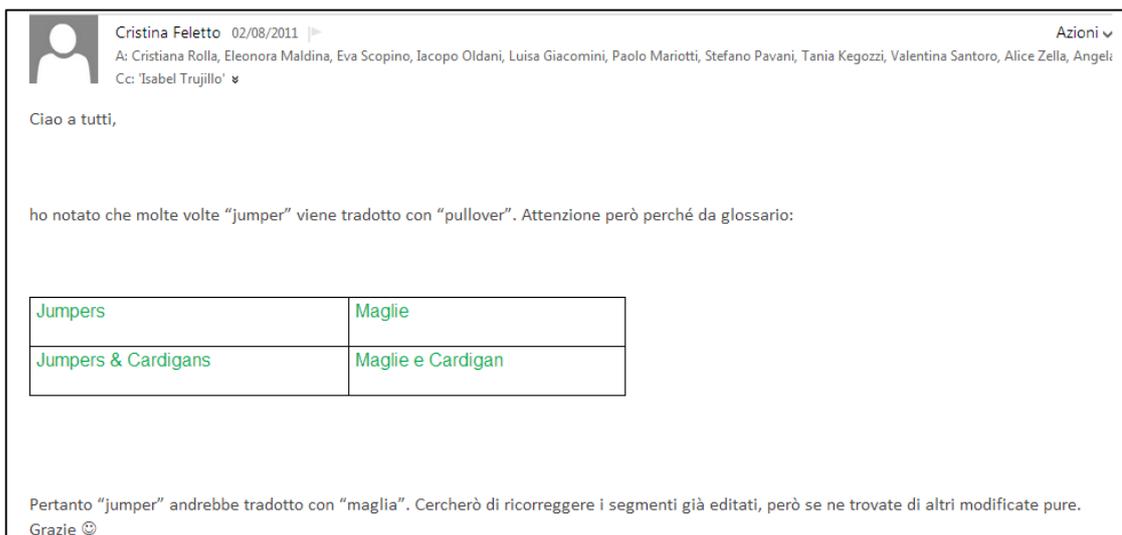


Fig. 16: Una conversazione per email relativa alla traduzione del termine "jumper"

Solitamente i glossari vengono condivisi in formato Excel, ma nei casi in cui l'agenzia richieda da parte dei traduttori l'utilizzo di un CAT Tool, allora il glossario verrà fornito in un formato specifico adatto ad essere importato all'interno del software e ad essere visualizzato, adoperato e modificato nel corso del lavoro. In Trados, lo strumento di gestione della terminologia è denominato Multiterm. Come in diversi altri CAT Tools esso può essere utilizzato *stand-alone* o integrato con la memoria per fare in modo che i due strumenti si interfaccino e si scambino informazioni in corso d'opera. Attivando la funzione "Active Term Recognition" per il riconoscimento della terminologia presente nel testo da tradurre, il CAT Tool scansionerà ogni segmento di testo confrontandolo con uno o più glossari selezionati per il progetto corrente; quando troverà una corrispondenza, visualizzerà il risultato della ricerca all'interno di una finestra apposita, dalla quale sarà possibile inserire il termine direttamente nel testo. Allo stesso modo, sarà possibile inserire nuova terminologia mentre si traduce. Basterà selezionare prima il termine nella lingua di partenza, poi quello nella lingua di arrivo, per creare una nuova voce di glossario che potrà essere utilizzata immediatamente e ampliata con informazioni aggiuntive se necessario. L'immagine sotto riportata mostra un glossario creato tramite Multiterm. Si distinguono tre finestre principali, quella in alto a sinistra mostra tutte le voci del glossario, quella sottostante i glossari aperti al momento e i preferiti, e quella principale mostra i particolari della voce selezionata.

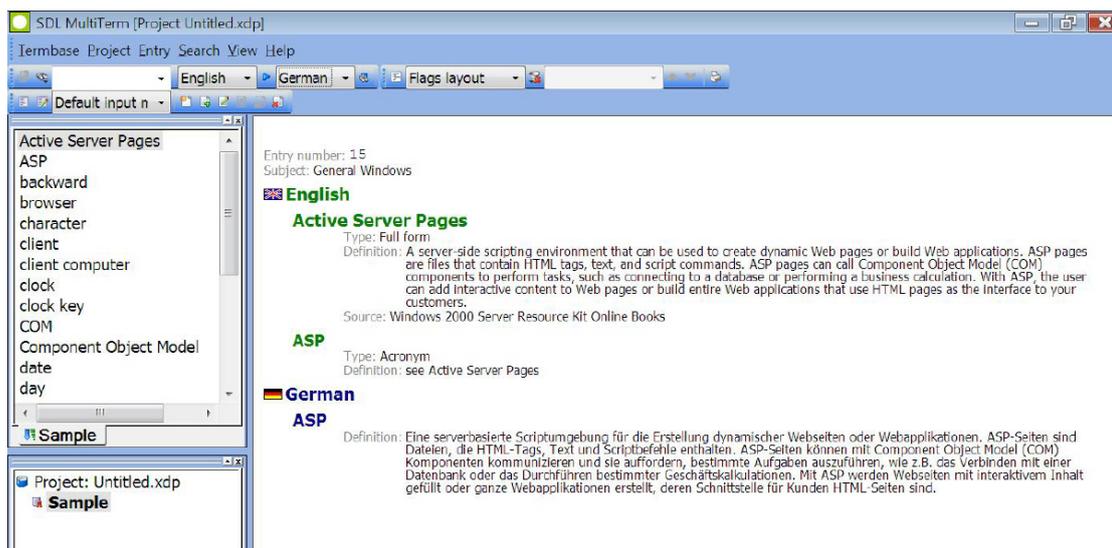


Fig. 17: Un glossario in SDL Multiterm

Nell'esempio sopra riportato il glossario è bilingue, inglese-tedesco. Con Multiterm, tuttavia, così come con altri strumenti consimili, è anche possibile cambiare l'ordine delle due lingue (nell'esempio, diventerebbe tedesco-inglese), nonché creare glossari multilingui, dai quali selezionare in base al bisogno le lingue da visualizzare, sia quella da cui ricercare che quella verso cui si dovrà tradurre. Come si può vedere dall'immagine, è inoltre possibile corredare la voce di glossario con diversi tipi di informazione, come il tipo di voce (forma intera o acronimo nell'esempio), una definizione, la fonte dalla quale questa è stata presa, ma anche immagini esemplificative, annotazioni per il traduttore, collegamenti ad altre voci di glossario o a pagine esterne. La voce può anche essere modificata una volta creata, grazie al comando "Edit" che in tempo reale consente di correggere o aggiungere informazioni. Quando invece il glossario verrà consultato durante la traduzione, esso sarà visualizzato direttamente all'interno dell'ambiente di lavoro, in un riquadro all'interno del Workbench, dal quale non sarà possibile compiere modifiche, ma solo aggiungere nuovi termini *on the fly*, come già accennato. Il glossario completato potrà essere esportato e salvato in un formato di interscambio che lo renderà condivisibile con altri colleghi o utilizzabile tramite altri software che ne supportino il formato. Esso potrà inoltre essere salvato sul computer del traduttore o su dispositivi esterni senza occupare grandissimo spazio a prescindere dall'ampiezza della raccolta terminologica.

### 2.3.2. I terminology extraction tools

Un software a parte, di altrettanta utilità, è Multiterm Extract, la cui finalità è quella di estrarre terminologia dal confronto tra documenti monolingui o bilingui. Il metodo utilizzato si basa su un calcolo statistico atto a determinare la frequenza della comparsa di termini candidati all'interno del testo, allo stesso modo di come gli strumenti di analisi dei corpora calcolano i *token*, e di come il *Concordancer* recupera tutte le occorrenze di un dato termine nel suo contesto, come si è visto al paragrafo 1.4 del capitolo precedente. Multiterm Extract analizzerà i testi e creerà delle liste di termini ricorsivi in essi, e nel caso della ricerca in documenti bilingui, affiancherà la probabile traduzione di tali termini, individuata tramite il calcolo del numero delle occorrenze e della sua comparsa in concomitanza all'utilizzo del termine nella lingua di partenza. Il traduttore individuato dovrà comunque essere confermato dal traduttore per maggiore sicurezza. Una volta create le liste e confermate le traduzioni dei termini, i risultati potranno essere esportati nel formato di MultiTerm, XML, e importati all'interno di un glossario. In tal modo sarà possibile valutare la qualità della terminologia di cui si è già in possesso confrontandola con quella utilizzata in documenti che ne mostrino l'utilizzo reale, o al contrario si potrà anche verificare il corretto utilizzo della terminologia all'interno di un testo tradotto confrontandolo con il database creato. Questo strumento può infine essere utilizzato alla pari di un *corpus-analysis tool*, per studiare testi appartenenti allo stesso ambito e rilevarne la terminologia e i costrutti più frequenti, quindi in ambito didattico o per lo studio personale del traduttore che intenda specializzarsi in un determinato settore. Strumenti simili a Multiterm Extract sono offerti da altri CAT Tools, tra i quali la suite MemoQ (<http://kilgray.com/products/memoq>), Wordfast nelle sue articolazioni Pro, Classic e Anywhere<sup>13</sup>, ma esistono anche numerosi sistemi gratuiti online<sup>14</sup>.

---

<sup>13</sup> Si veda <http://www.freetm.com/> per visualizzare la versione open source di Wordfast.

<sup>14</sup> Una ampia carrellata viene offerta in <http://recremisi.blogspot.it/p/online-terminology-tools.html>.

## 2.4 Strumenti per la localizzazione e il project management

### 2.4.1. Strumenti per la localizzazione

Per presentare una panoramica quanto più possibile completa degli strumenti disponibili per l'industria della traduzione, è necessario accennare anche ai software che vengono in sostegno dei progetti di localizzazione e a quelli adoperati per il project management. In generale si tratta di strumenti molto articolati, che devono dare risposta a una molteplicità di esigenze dovute alla complessità di queste attività. Della localizzazione si è già discusso nel primo capitolo, e si è già descritta l'ampiezza di tale processo, che implica un lavoro su più livelli per rendere un prodotto, principalmente un software ma anche un sito web o un videogioco, perfettamente funzionale per il paese al cui mercato esso è destinato, e coerente con la sua cultura, le sue consuetudini, le sue normative. Poiché il termine "localizzazione" si riferisce nella maggior parte dei casi a prodotti tecnologici e digitali, un'importante funzione che i software dedicati devono svolgere è quella di tutelare il codice di programmazione per non alterarlo, estrapolando il testo da tradurre allo stesso modo in cui il TagEditor protegge tramite l'utilizzo dei *placeable* la formattazione e le altre parti del testo da non modificare. Un altro compito importante svolto dai software per la localizzazione è quello di adattare il testo tradotto, spesso più lungo di quello originale, alla videata nella quale esso apparirà. Sarà possibile per esempio ridurre leggermente l'immagine o modificare la visualizzazione del testo, grazie alla fase detta di "pseudo-traduzione", durante la quale il testo originario viene preparato ad essere tradotto aggiungendo alcuni parametri di lunghezza delle stringhe di testo o la possibilità di supportare e visualizzare correttamente caratteri speciali come le lettere accentate, anche se esse non sono previste nel linguaggio di partenza. Nel caso di differenze considerevoli, il software stesso in alcuni casi provvederà a segnare un'allerta da rivolgere al team di *software engineering*.

I sussidi per la localizzazione ricorrono ovviamente anche agli strumenti utilizzati per i progetti di traduzione, quindi le memorie, i glossari e gli strumenti di QA. Essi devono inoltre agevolare il lavoro in team, fornendo a tutte le figure professionali coinvolte nel progetto delle basi comuni da adottare per scambiarsi documenti e risorse. I traduttori lavorano a stretto contatto con gli sviluppatori del software al fine di accelerare quanto più possibile il lancio delle versioni localizzate. I traduttori possono prospettare ai programmatori determinati elementi che potrebbero costituire un ostacolo in fase di

traduzione, mentre dal canto loro i programmatori possono fornire ai traduttori supporto tecnico e informazioni sulla struttura del prodotto. I project manager fanno da *trait d'union* tra di essi e compilano database terminologici e materiale di supporto. Per poter lavorare veramente in sintonia, questi professionisti necessitano tuttavia di condividere un ambiente di lavoro che sarà rappresentato dalla piattaforma di localizzazione.

Sono molti i software per la localizzazione. Tra i più noti Passolo, prodotto da SDL, ed Alchemy Catalyst.

#### **2.4.2. Strumenti per il *project management***

Per passare ai programmi utilizzati per la gestione dei progetti, è necessario soffermarsi un istante sulla complessa articolazione di attività e compiti che possono essere inclusi all'interno del *project management*. I Project Manager sono figure fondamentali all'interno delle agenzie di traduzione. Essi svolgono una importante funzione di mediazione tra il cliente e il team di traduttore, e di comunicazione tra i vari professionisti che con compiti diversi contribuiscono al progetto di traduzione. Le grandi agenzie di traduzione contano su un project manager per ogni grosso progetto di traduzione, o al massimo per due allo stesso tempo, se non si tratta di progetti molto grandi e impegnativi. Egli dovrà occuparsi di seguire i lavori dal momento in cui il cliente affida l'incarico all'agenzia, quindi nella fase di trattativa dei costi e delle tempistiche di consegna, fino al completamento totale della fase di editing e proofreading, quando avviene la consegna del progetto tradotto. I compiti del project manager sono quindi essenzialmente di natura economica, gestionale, di controllo e mediazione. Deve sapere gestire grandi quantità di testi e materiali di riferimento, mantenerli sempre aggiornati e garantirne la correttezza e l'affidabilità; deve saper rispondere in maniera certa ed esaustiva ai quesiti dei traduttori, saper trovare le giuste soluzioni con il cliente laddove si presentassero delle difficoltà e assicurarsi che i suoi collaboratori rispettino le linee guida e le scadenze date loro; deve inoltre verificare che il progetto rientri nel budget stabilito dall'agenzia sulla base del prezzo concordato con il cliente, che vi siano sufficienti traduttori operanti nel progetto e che gli editor svolgano un lavoro di alta qualità per garantire un risultato finale ottimale.

L'immagine seguente sintetizza in modo chiaro e dettagliato le varie fasi della gestione dei progetti di traduzione, dalla pianificazione iniziale alla chiusura e fatturazione:

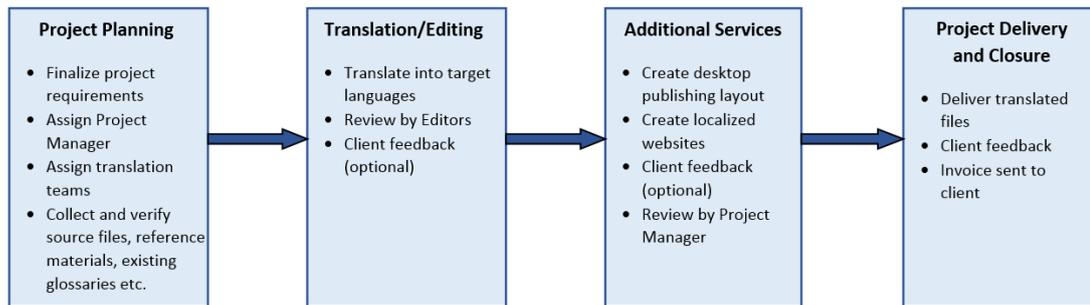


Fig. 18: Le varie fasi di un progetto di traduzione (Esselink, 2000)

Per far fronte a tale complessità di mansioni, poter contare su strumenti che semplifichino il lavoro e contribuiscano alla verifica dei parametri essenziali di qualità del prodotto finale è di indubbia utilità e valore per i project manager. Tali software seguono il processo passo dopo passo, si interfacciano con gli altri programmi utilizzati per importare da essi i risultati delle analisi compiute sui testi ed utilizzarli per finalità di verifica e controllo, consentono di conteggiare i compensi per le varie figure professionali coinvolte (traduttori, revisori, addetti al controllo terminologico ecc.), e di calcolare il carico di lavoro giornaliero richiesto per terminare il progetto entro la data di consegna stabilita. Oltre ai software specifici necessari alle sue mansioni, il project manager dovrà ovviamente conoscere e saper utilizzare i CAT Tools e le risorse utilizzate dai traduttori, non solo per rivedere la traduzione all'interno dell'ambiente in cui è stata compiuta e poter apportare eventuali modifiche consultando e aggiornando in tempo reale memorie e glossari, ma anche perché eventuali problemi di natura tecnica con tali programmi saranno indirizzati prima a lui e solo dopo all'esperto incaricato della manutenzione dei software aziendali.

#### 2.4.3. Le suite SDL Trados

La suite dei prodotti Trados 2007 includeva una piattaforma denominata Synergy dedicata ai project manager, mentre nelle versioni successive questa è stata inserita nell'unico ambiente detto Studio, alla voce "Projects" del menu a sinistra.

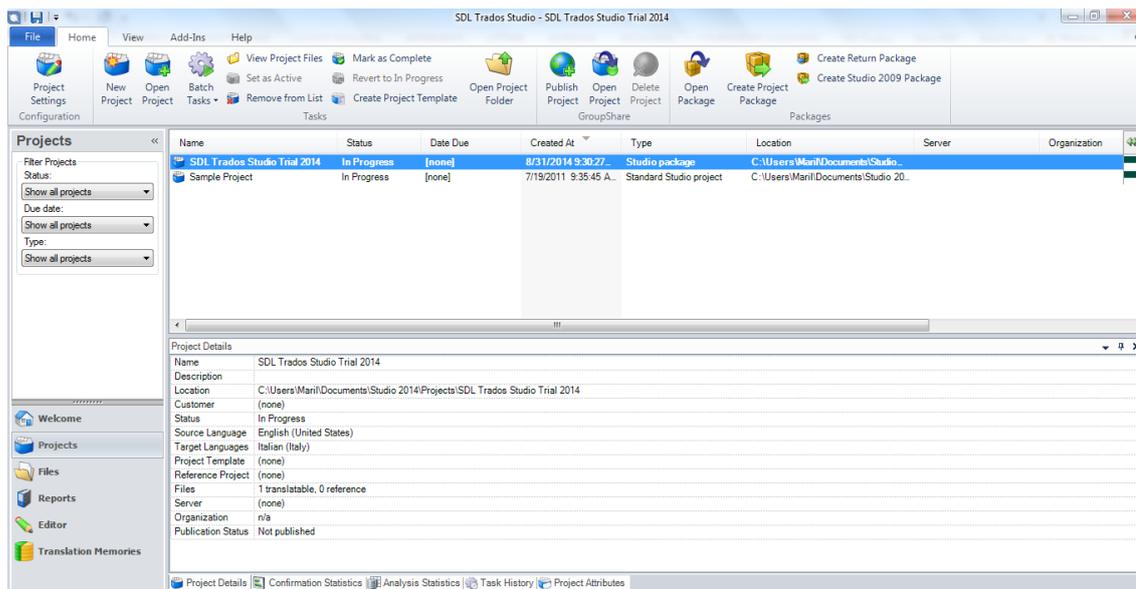


Fig. 19: La piattaforma di Project in Trados Studio

Da questa videata è possibile verificare i progetti sui quali si sta lavorando. Cliccando su ognuno di essi, si accederà alle informazioni relative, come l'insieme dei file, i tempi di consegna, il grado di avanzamento del lavoro espresso in percentuale ecc. Da questa schermata il project manager potrà:

- Creare un nuovo progetto
- Distribuire il lavoro ai membri del team creando i così detti “project packages”
- Tenere aggiornate le linee guida relative al progetto
- Aggiungere nuovi file a progetti preesistenti
- Aggiungere TM o glossari a progetti preesistenti
- Convertire i file tradotti dal formato di lavoro nel formato originario
- Segnare i progetti come completi
- Inviare il tutto al cliente
- Condividere col team il feedback del cliente

I *project packages* non sono altro che “pacchetti”, o per meglio dire cartelle compresse, contenenti i documenti da tradurre e i materiali di riferimento, ovvero TM e glossari nei formati necessari ad essere importati nel Workbench del traduttore e utilizzati, ampliati e

aggiornati durante il suo lavoro. Una volta terminato il batch di traduzioni affidategli, il traduttore dovrà avere cura di ritornare debitamente aggiornati tutti i documenti ricevuti, più i file tradotti ed eventuali altri file prodotti durante il suo lavoro. Per esempio, laddove egli inserisca dei commenti nel testo, essi verranno salvati su un file esterno, che se non verrà inserito nel package renderà impossibile per il project manager visualizzare tali commenti. Le TM e i glossari utilizzati dovranno essere aggiornati e salvati nel formato in cui sono stati ricevuti, e nella cartella dovranno essere presenti il file originario, quello tradotto e il file bilingue prodotto dal CAT Tool (nel formato adatto ad essere visualizzato e modificato nel TagEditor). La correttezza nell'invio e nella ricezione di tali file è vitale per la qualità e la gestione ottimale del progetto, soprattutto quando si gestiscono quantità di documenti non indifferenti.

Nel nuovo Trados Studio una sezione a parte è dedicata ai report, ovvero alle analisi dei testi originali, finalizzate a conoscere la loro composizione a partire da un confronto con le memorie e i glossari esistenti. Queste analisi infatti riveleranno il numero di segmenti già presenti in memoria, sia come *exact matching* che come *fuzzy matching*, il numero di ripetizioni interne e la percentuale di testo che invece dovrà essere tradotta ex novo. Verrà rivelata anche la percentuale di terminologia già nota e presente nei glossari; in tale maniera il traduttore potrà comprendere il tipo di impegno che quel dato progetto comporterà, e il project manager potrà stabilire anche costi e tempistiche del lavoro da poter comunicare al cliente.

Un'opzione molto comoda per i project manager è quella di filtrare i progetti in base alla data di consegna, in modo da avere sempre presenti quelli più urgenti e dalle scadenze più ravvicinate. Diverse altre opzioni specifiche sono disponibili, quello che è comunque più importante sottolineare è che Trados Studio, così come gli altri CAT Tools nelle loro versioni più moderne, consentono il beneficio dell'integrazione delle funzioni in un unico ambiente di lavoro, a differenza dei diversi applicativi che occorreva avviare separatamente nelle versioni precedenti. Avere tutti gli strumenti racchiusi in un'unica soluzione conferisce snellezza e velocità al lavoro, consente di passare facilmente dal lavoro di traduzione a quello di revisione e gestione dei progetti o organizzazione dei materiali, a vantaggio della qualità del risultato finale.

## 2.5 *Le risorse disponibili su internet*

Dedichiamo l'ultima parte di questo capitolo alla descrizione di alcuni utili strumenti e risorse che, seppur non essendo classificabili come CAT Tools, vengono quotidianamente utilizzati dai traduttori come supporto linguistico o stilistico e rappresentano una fonte preziosa di informazioni e materiali facilmente ottenibili nella maggior parte dei casi anche gratuitamente. Si tratta dell'infinita varietà di risorse disponibili su internet, che ogni buon traduttore conosce e ricerca per trovare risposta a diversi tipi di esigenze (Fulford e Granell Zafra, 2004: 39). Non verranno trattati in questo luogo i vari strumenti di traduzione automatica, perché ad essi verrà dedicato ampio spazio nella Seconda Parte di questa tesi. Ci si concentrerà piuttosto sulle tante pagine di internet dedicate principalmente a due aspetti, quello della ricerca terminologica e quello del confronto tra professionisti del settore.

### **2.5.1. Le pagine dedicate alla terminologia**

Su internet sono disponibili per i traduttori diversi siti dai quali ottenere TM<sup>15</sup> e glossari, nonché corpora monolingui o plurilingui. Tra i più famosi, il sito della Direzione generale della Traduzione dell'Unione Europea<sup>16</sup> che dal 2007 ha reso disponibili a tutti le proprie memorie di traduzione nelle 23 lingue ufficiali dei suoi stati membri; anche la Commissione Europea fornisce gratuitamente TM e materiali linguistici raccolti a partire dal 2004<sup>17</sup>. Le TM sono fornite nel formato di interscambio (TMX – Translation Memory Exchange) in modo da poter essere importate da qualsiasi CAT Tool. Pur con una preponderanza del linguaggio giuridico/legislativo, sono presenti anche numerosi altri ambiti linguistici tecnici, come quello dei brevetti, dell'agricoltura, dell'economia e di tutte le attività svolte o seguite dai vari uffici delle istituzioni europee.

---

<sup>15</sup> Si veda su quest'argomento Levitt, 2003

<sup>16</sup> <http://open-data.europa.eu/it/data/dataset/dgt-translation-memory>

<sup>17</sup> <http://ipsc.jrc.ec.europa.eu/index.php?id=197>

The screenshot shows the 'European Union Open Data Portal' interface. At the top, there is a navigation bar with 'Data', 'Applications', 'Linked Data', and 'About' tabs. The main content area is titled 'DGT-Translation Memory'. It includes a 'Publisher' section with the 'Directorate-General for Translation' link, a 'Description' section with text about the translation memory in 23 languages, a 'Licence' section with 'Europa Legal Notice', a 'Keywords' section with terms like 'corpus' and 'translation-memory', and a 'Resources' section with download buttons for a detailed description (PDF) and two volumes (ZIP).

Fig. 20: La pagina della Direzione Generale della Traduzione dalla quale è possibile scaricare le TM

Di siti che offrono dizionari e glossari ve ne sono infiniti per le diverse combinazioni linguistiche. Alcuni siti ne raccolgono diversi e li catalogano per argomento e per lingua, in altri casi invece vi sono pagine espressamente dedicate ad argomenti più o meno specifici. Al traduttore alla ricerca di materiali per un determinato settore sarà sufficiente consultare i siti dedicati ai traduttori, di cui si parlerà più avanti, o digitare sul motore di ricerca la combinazione linguistica e l'argomento per trovare moltissime pagine e risorse utili. Tra i vari siti da cui attingere dizionari e glossari si possono annoverare [www.lexicool.com](http://www.lexicool.com), [www.gotranslators.com](http://www.gotranslators.com), [www.dixon.es](http://www.dixon.es) (con pagine per diverse lingue); tra le pagine specifiche per argomento possiamo portare ad esempio alcune dedicate alle traduzioni giuridiche, come [www.justranslation.com](http://www.justranslation.com) e [www.traduttorelegale.it](http://www.traduttorelegale.it).

Laddove invece il traduttore desiderasse ricercare un termine specifico, avrà a disposizione innumerevoli possibilità. Prima di tutto, una semplice indagine tramite motore di ricerca potrà chiarire alcuni dubbi, come ad esempio se l'espressione adottata nella traduzione sia realmente utilizzata. Il numero delle occorrenze che di quella espressione verranno recuperate dal motore di ricerca sarà indicativo del suo reale utilizzo, o in alcuni casi inserendo il termine parzialmente incorretto sarà il motore di ricerca stesso o le pagine che verranno visualizzate a suggerire l'espressione corretta. In

molti casi, soprattutto per le traduzioni dall'inglese, si potrebbe scoprire che il termine originario viene utilizzato anche nella lingua di arrivo, specialmente se questa è l'italiano, una lingua che predilige i prestiti linguistici. Un trucco molto utile in diverse circostanze consiste nell'utilizzare Wikipedia, la celebre enciclopedia multilingue online. Per ogni voce, essa riporta anche le altre lingue nelle quali il termine ricercato è disponibile (in basso a sinistra, evidenziato in rosso). Basterà cliccare sulla lingua verso la quale si sta traducendo per visualizzare la pagina con la traduzione del termine:

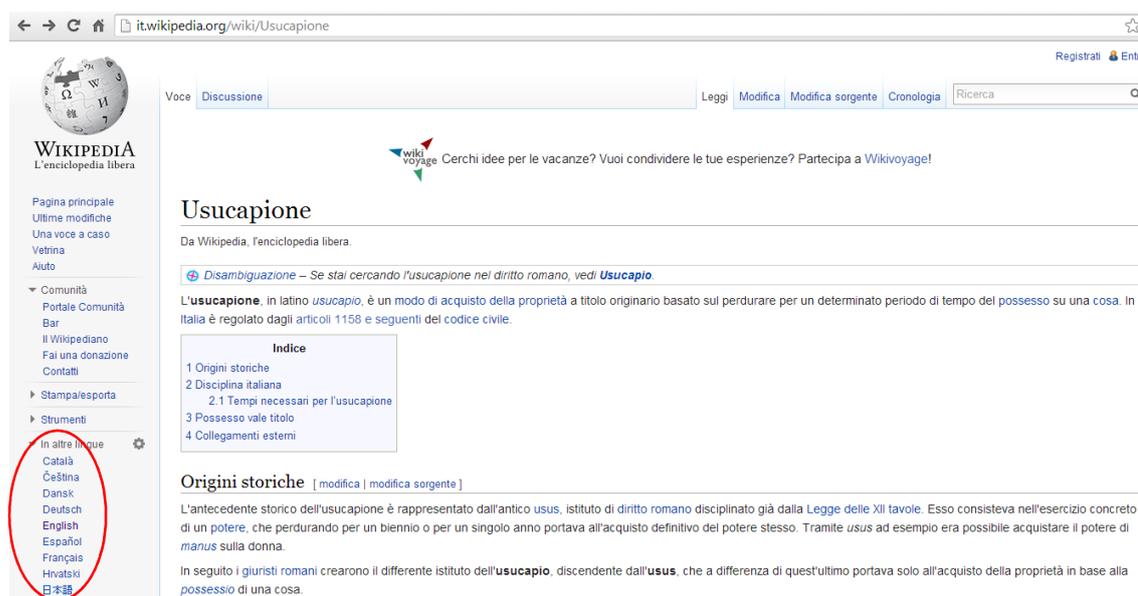


Fig. 21: Una pagina di Wikipedia, nell'ovale rosso le altre lingue disponibili

Una risorsa gratuita davvero preziosa viene fornita dall'agenzia Translated.net. Essa consente agli utenti di ricercare termini ed espressioni all'interno delle proprie memorie di traduzione selezionabili per lingua di partenza e di arrivo nonché per argomento. Come si vede dall'immagine sottostante, viene fornita una traduzione automatica del termine ricercato, seguita dai contributi umani presenti in memoria, con dei dettagli che possono tornare utili al fruitore, come l'argomento del testo in cui l'espressione è utilizzata, informazione che può rivelarsi molto importante nel caso di termini fraintendibili o con traduzioni diverse in base al loro contesto di utilizzo. È anche possibile scaricare l'intera TM in formato TMX o contribuire al servizio condividendo una propria TM.

The screenshot shows the MyMemory website interface. At the top, there is a search bar containing the text "to all intents and purposes". Below the search bar, there are options for language pairs (English to Italian) and a search button. The main content area is divided into two sections: "Traduzione automatica" and "Contributi umani".

Traduzione automatica		
Inglese	Italiano	Dettagli
to all intents and purposes	a tutti gli effetti	Da: Traduzione automatica Suggerisci una traduzione migliore Qualità: ☆☆☆☆☆ Sarai il primo a votare

Contributi umani		
Inglese	Italiano	Dettagli
It is, to all intents and purposes. <a href="http://www.europarl.europa.eu/">http://www.europarl.europa.eu/</a>	In pratica di questo si tratta. <a href="http://www.europarl.europa.eu/">http://www.europarl.europa.eu/</a>	Ultimo aggiornamento: 2012-03-23 Argomento: Scienze Sociali Frequenza d'uso: 3 Qualità: ☆☆☆☆☆ Sarai il primo a votare
This manoeuvre has to all intents and purposes been defeated. <a href="http://www.europarl.europa.eu/">http://www.europarl.europa.eu/</a>	Questa manovra è stata bocciata in tutti gli scopi e intenti. <a href="http://www.europarl.europa.eu/">http://www.europarl.europa.eu/</a>	Ultimo aggiornamento: 2012-02-29 Argomento: Scienze Sociali Frequenza d'uso: 1 Qualità: ☆☆☆☆☆ Sarai il primo a votare
To all intents and purposes these people have no access to health care. <a href="http://www.europarl.europa.eu/">http://www.europarl.europa.eu/</a>	In nessun caso queste persone hanno accesso all'assistenza sanitaria. <a href="http://www.europarl.europa.eu/">http://www.europarl.europa.eu/</a>	Ultimo aggiornamento: 2012-02-29 Argomento: Scienze Sociali Frequenza d'uso: 1 Qualità: ☆☆☆☆☆ Sarai il primo a votare

Fig. 22: MyMemory, la risorsa messa a disposizione da Translated.net

### 2.5.2. Le pagine dedicate ai traduttori

Come questa prima tipologia di risorse, anche quelle dedicate allo scambio e alla condivisione professionale sono molto abbondanti e molto utili, fornendo diversi servizi e vantaggi ai traduttori. Esiste una grande varietà di siti dedicati, alcuni più specifici per dati argomenti, altri più ampi e differenziati al loro interno. I più famosi sono sicuramente Proz (www.proz.com) e Translator's Café (www.translatorscafe.com), portali web multilingui dedicati a tutti i professionisti della traduzione, nei quali i traduttori possono ricercare lavori o proporsi ai potenziali clienti, possono confrontarsi tra colleghi su diversi argomenti, o ricercare e condividere informazioni o materiali utili per il proprio lavoro. Inoltre questi portali sono anche generalmente frequentati da studenti o traduttori alle prime armi al fine di trovare consigli, informazioni e indicazioni da professionisti più esperti (McDonough, 2007: 793). L'immagine seguente mostra una videata di Proz.com, in particolare una discussione all'interno del forum tematico offerto dal sito. Le discussioni in questo forum sono divise per argomento, mentre il forum dedicato al confronto su questioni terminologiche (www.proz.com/kudoz/) è diviso per combinazione linguistica, e all'inizio di ogni nuova discussione viene anche specificato l'ambito linguistico di pertinenza.

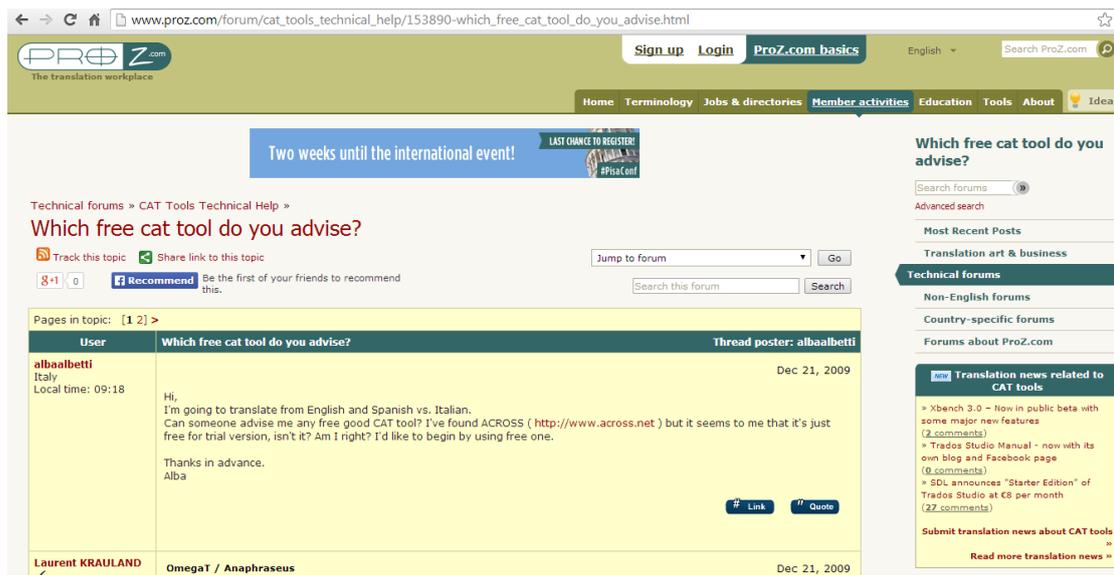


Fig. 23: Una discussione sul forum di ProZ.com

Gli argomenti su cui è possibile confrontarsi sono molteplici. Oltre alla terminologia vi sono discussioni dedicate alle dinamiche e le problematiche professionali quali la retribuzione e le questioni fiscali, o l'attendibilità di agenzie con le quali si ha intenzione di avviare una collaborazione e gli eventuali episodi di frode che purtroppo sono numerosi anche in questo ambito, o ancora corsi, master, seminari e *webinar* (seminari in audio-videoconferenza) per la propria formazione e crescita professionale, ecc. Un servizio newsletter informa gli iscritti delle nuove offerte disponibili sul sito, come sconti su CAT Tools disponibili per i soci, corsi gratuiti online, nuovo materiale linguistico disponibile per le combinazioni linguistiche dichiarate dall'utente ecc.

I siti per traduttori sono il mezzo principale attraverso il quale la domanda e l'offerta si incontrano. I clienti privati pubblicano le loro richieste di lavoro, le agenzie di traduzione ricercano collaboratori per ampliare il proprio organico o il proprio bacino di freelancer, e i traduttori presentano il proprio profilo professionale corredato di informazioni e contatti e possono proporsi per le offerte di lavoro presenti. La figura successiva mostra il sito Translatorscafe.com e in particolare la pagina delle offerte e richieste di lavoro, che è possibile filtrare per combinazione linguistica. Nei riquadri laterali si possono distinguere offerte pubblicitarie e link a pagine correlate quali quelle dedicate alla ricerca di freelancer o agenzie, alla pubblicazione di un'offerta ecc.

www.translatorscafe.com/cafe/SearchJobs.asp

Lavori | TCTerms | Forum | Sondaggi | Articoli | Membri | Glossari

Tutti: 0.000 Selezionato: 0.000

**May 31, 2014: Lavori di traduzione recenti**

Ci sono 200 lavori di traduzione recenti in **tutte le combinazioni linguistiche.**

Pag.: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

Thanks for the feedback! [Back](#)  
 We'll review this ad to improve your experience in the future.  
 Help us show you better ads by updating your [ads settings](#).

#168182	Need translator for Spanish to English lanugae pair(Medical) Traduzione	Spagnolo>Inglese	09:34 GMT
#168181	Thai-English translation Traduzione	Thailandese>Inglese	09:24 GMT
#168180	Reiured translation for Finnish to English Traduzione	Finlandese>Inglese	08:27 GMT
#168179	Urgent opening for Tamil speaker at Noida based company Altro	Inglese>Tamil	07:01 GMT
#168178	Urgent opening for Kannada speaker at Noida based company Altro	Inglese>Kannada	06:57 GMT
#168177	English to Urdu Translation of Legal content. Traduzione	Inglese>Urdu	06:19 GMT

Noncredit Courses and Certificates  
 Translation and Interpreting  
 Enroll Today for Summer ▶

Funzioni e produttività straordinarie.  
 35 giorni  
 Rimborso di € 150\*  
 clicca per i prezzi  
 Stampante multifunzione  
 fino a 4 colori  
 Xerox® WorkCentre® 6605

Fig. 24: Translators' Café, la pagina delle offerte di lavoro

In ultimo, è giusto accennare anche alle possibilità offerte dai social network, in particolare Facebook e LinkedIn. Quest'ultimo è un social di natura prevalentemente professionale, nato con l'intento di mettere in contatto gli utenti sulla base dei loro profili professionali ed esperienze lavorative, creando una rete di collegamenti rilevanti per il proprio lavoro. I contatti di LinkedIn sono principalmente colleghi di università o di lavoro, persone con le quali si è condivisa un'esperienza formativa o lavorativa, o persone operanti nello stesso settore che possono rappresentare potenziali occasioni future di collaborazione professionale. L'utente compila il proprio profilo inserendo i dati relativi alla propria formazione universitaria e post-laurea, le proprie esperienze lavorative e gli enti, le società o i privati presso i quali si sono maturate tali esperienze, le proprie competenze e conoscenze. I contatti dell'utente possono lasciare un feedback sul suo profilo, confermando le competenze riportate sulla pagina, o raccomandando l'utente per la sua professionalità in modo che altri potenziali clienti possano rivolgersi a lui. In poche parole, LinkedIn è una perfetta vetrina in cui un professionista può mostrare e offrire le proprie esperienze e conoscenze ed avere qualcuno che confermi e garantisca le competenze che egli sostiene di possedere.

Facebook è il social network per antonomasia, nato allo scopo di mettere in contatto persone con le quali si è condiviso qualcosa nella propria vita. Tuttavia, al di là del suo uso prettamente sociale e relazionale, esso è anche diventato un "luogo" di incontro, confronto e interscambio tra professionisti. Basta fare una rapida ricerca per

vedere quante pagine sono dedicate alla traduzione, dalle pagine generiche a quelle che pubblicizzano corsi di formazione, a quelle in cui i traduttori possono discutere le loro problematiche professionali, condividere offerte di lavoro o chiedere un aiuto pratico come nell'immagine seguente:



Fig. 25: Una discussione su un gruppo Facebook

L'utilità di Facebook è data da un lato dal suo grande e crescente bacino di utenza, dall'altro dal fatto che molti dei suoi utenti sono collegati alla pagina quasi sempre anche mentre fanno altro, e possono quindi visualizzare e rispondere a discussioni quasi in tempo reale o comunque in tempi molto brevi.

Da questa rapida panoramica sulle molteplici possibilità offerte in rete appare chiaro come conoscerle ed utilizzarle in maniera accorta e consapevole sia senza dubbio una reale opportunità per i traduttori. La ricchezza e la molteplicità di tale strumento gioca indubbiamente a vantaggio dei professionisti che potranno ricercare informazioni e risorse rapidamente, semplicemente e in uno spazio virtuale grande come il web.

### *3.1 Dentro il processo traduttivo*

Dopo aver tracciato un esauriente quadro sui CAT Tools, sulle loro funzionalità, la loro classificazione e le teorie linguistiche a partire dalle quali essi sono stati sviluppati, il presente capitolo cercherà di addentrarsi nel processo traduttivo per comprendere come l'introduzione della tecnologia ne modifichi alcune dinamiche. Occorrerà prima di tutto chiarire alcuni concetti a monte, ovvero cosa si intende per “processo traduttivo”, quali operazioni esso racchiude e che grado di generalità possiamo attribuire ad esso, tenendo in conto la differenza di approcci, strategie e procedure che ogni traduttore decide, più o meno consapevolmente, di mettere in atto.

#### **3.1.1. La traduttologia cognitiva**

Essendo a pieno titolo considerabile come un atto comunicativo, e per di più particolarmente complesso secondo quanto spiegato da Neubert<sup>18</sup> (1997: 6-7), la traduzione è in prima istanza il frutto di un processo cognitivo che avviene nella mente del traduttore e che scaturisce dalla sua formazione, esperienza e in parte anche dalla sua stessa personalità. Questo aspetto è stato approfondito dall'approccio cognitivista alla traduzione, con lo sviluppo negli ultimi anni di una “traduttologia cognitiva” (Martín 2007; Amigo Extremera, 2013), che ha messo in luce la complessità del processo traduttivo dal punto di vista dei meccanismi mentali, consapevoli e non, che esso implica. La complessità risiede nell'articolazione di tale processo in diverse fasi, nella varietà di competenze che esso richiede e di risorse che mette in gioco, ma anche nella sua stessa natura di processo mentale prima che pratico, che implica l'impossibilità di osservare tale processo in maniera diretta. Per studiare questo aspetto della traduzione si è fatto ricorso alle neuroscienze (Salmon, 2007 e 2005), alla linguistica cognitiva e persino alla psicologia e alla sociologia, per cercare di comprendere il funzionamento della memoria,

---

<sup>18</sup> In Danks, Shreve, Fountain, Mc Beath (1997)

la capacità di lettura e analisi del testo, il meccanismo di suddivisione del testo in porzioni minime, i criteri che regolano la ricerca e l'applicazione di strategie, la trasposizione prima mentale e poi scritta del messaggio da una lingua a un'altra, nonché i condizionamenti interni ed esterni che il traduttore subisce, la sua personale percezione del lavoro che svolge (Kiraly, in Danks *et al.*, 1997), la gestione dello stress e diversi altri aspetti non convenzionalmente studiati soprattutto nei primi anni di sviluppo dei Translation Studies, ma oggi considerati estremamente importanti per comprendere appieno il processo traduttivo.

### 3.1.2. Il modello di Kiraly

L'immagine seguente è il modello psicolinguistico sviluppato da Kiraly (1995: 101).

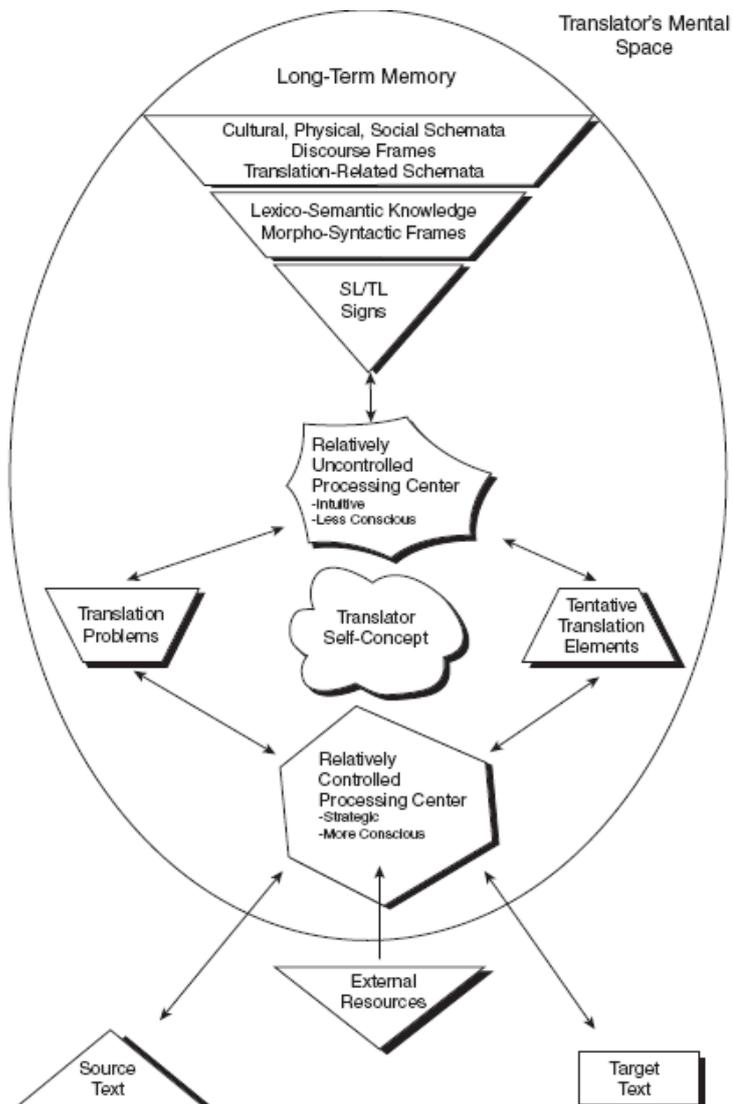


Fig. 25: Il modello psicolinguistico di Kiraly

Esso mostra il processo traduttivo come un'attività cognitiva. L'ovale racchiude tutto ciò che avviene o è immagazzinato all'interno della mente del traduttore, gli elementi al di fuori di esso rappresentano invece gli stimoli che provengono dall'esterno, ovvero il testo originario, il testo target e le risorse utilizzate. Questi elementi vengono acquisiti ed elaborati da quella che Kiraly chiama "Relatively controlled Processing center" (Centrale di elaborazione relativamente controllata), la parte maggiormente consapevole della mente del traduttore che utilizza la memoria a breve termine e nello specifico la sua capacità operativa e di lavoro per determinare le migliori strategie e soluzioni. Nel confrontarsi con il testo da tradurre, ma come sottolinea Neubert (in Danks *et al.*, 1997) già nell'affrontare l'idea stessa di dover tradurre un dato testo per un dato pubblico, altri stimoli vengono forniti dall'interno, dal grande bagaglio di conoscenze e concetti precedentemente acquisiti e salvati nella memoria a lungo termine: si tratta di schemi culturali e sociali che il traduttore si è costruito ma anche delle conoscenze linguistiche e delle competenze professionali necessarie per comprendere il testo nella lingua originaria e riformularlo in quella di arrivo. Queste informazioni convergono nella "Relatively uncontrolled processing center" (Centrale di elaborazione relativamente incontrollata), nella quale avvengono i processi mentali meno consapevoli e più intuitivi. Le due centrali tuttavia non funzionano in maniera separata, ma si interfacciano continuamente al fine di produrre dei tentativi di traduzione e delle soluzioni ai problemi che sorgono in corso d'opera.

Rispetto ad altri modelli del processo traduttivo, il modello di Kiraly ha il vantaggio di essere molto semplice e di mostrare chiaramente i principali passaggi dell'elaborazione mentale compiuta dal traduttore. Esso potrebbe essere ampliato e particolareggiato con ulteriori elementi, come ad esempio riportando altri tipi di stimoli che provengono dall'esterno come le condizioni di lavoro ed altre forme di possibili condizionamenti, o maggiori informazioni sui "cassetti" della memoria che vengono esplorati al fine di trarre le informazioni necessarie al lavoro. Per gli scopi del presente capitolo, si cercherà adesso di ipotizzare come i CAT Tools o gli altri strumenti di sussidio alla traduzione possano trovare spazio all'interno di questo schema, per determinare in questo modo il tipo di funzione che essi svolgono e il modo in cui essi modificano il processo traduttivo. Lo scopo per il quale essi sono stati sviluppati è quello di agevolare il lavoro del traduttore svolgendo alcuni dei compiti più ripetitivi e meccanici che causano

dispendio di tempo ed energie. Progettati per venire in soccorso dei processi naturali messi in atto dalla mente del traduttore, essi introducono delle variabili operative che ne alterano la normale linearità, portando il traduttore a svolgere e addirittura percepire in maniera diversa il proprio lavoro. Questo tipo di studio, incentrato sugli effetti della tecnologia sui processi cognitivi in traduzione, non è ancora stato sviluppato in Italia, mentre ha interessato un certo numero studiosi in Europa, Canada ed America nell'arco degli ultimi dieci anni, corrispondenti all'affermazione sul mercato dei CAT Tools quali indispensabili strumenti di lavoro per i traduttori. In particolare, i principali studi si concentrano negli ultimi sei anni, e tentano di superare il punto di vista meramente funzionale che accompagna da sempre le memorie di traduzione e i Workbench. È indubbio infatti che essi siano fondamentalmente degli strumenti atti ad ottenere dei risultati apprezzabili dal punto di vista quali-quantitativo, ma è altrettanto vero che tradurre all'interno di un ambiente di lavoro determinato dall'utilizzo di un software implica un adattamento dell'intero processo a tale poliedrico strumento. Quindi, come già precedentemente affermato, la combinazione dei due campi d'indagine, gli studi cognitivi sul processo traduttivo da un lato e l'integrazione delle tecnologie in traduzione dall'altro, sembra poter fornire interessanti retroscena su come funziona la mente del traduttore (Alves e Hurtado Albir, 2010).

Tornando a riferirci al modello di Kiraly, essendo fornite da un supporto esterno alla mente del traduttore, le risorse messe a disposizione da un CAT Tool possono essere aggiunte alle altre risorse esterne che si trovano alla base dell'immagine. Tuttavia, le informazioni recuperate dalle memorie di traduzione possono essere al contempo paragonate a quelle recuperate dalla memoria a lungo termine, secondo quanto sostenuto da Christensen e Schjoldager (2011: 140):

*“A TM is to be considered a material artifact involved in the human translator’s process of organizing the functional skill of memorizing translation decisions made during earlier translations into cognitive functional skills. In other words, a TM constitutes a supplementary LTM<sup>19</sup> from which information can be retrieved. When groups of translators use a TM, they can be said to share cognitive resources via the artifact.”*

---

<sup>19</sup> Long Term Memory, ndr

Il recupero dei match dalla TM avviene in maniera automatica così come le informazioni della memoria a lungo termine (MLT) vengono recuperate in maniera inconsapevole, ed elaborate in maniera consapevole dalla memoria di lavoro. Allo stesso modo, anche il match, sia *fuzzy* che *exact*, che viene proposto automaticamente dalla TM deve sempre ottenere l'approvazione consapevole del traduttore e subire eventuali correzioni o rielaborazioni prima di essere considerato corretto e venire inserito nel testo. La stessa considerazione è valida anche per i termini o le espressioni estratti dai glossari di riferimento, non solo per le stringhe di testo salvate nella TM. Quando è attiva la funzione di "Active Term Recognition", il software recupererà automaticamente dal glossario i termini trovati all'interno del segmento di testo che si sta traducendo, e ne fornirà il corrispettivo nella lingua di arrivo.

Nello schema di Kiraly è presente un elemento di congiunzione tra le due centrali, che rappresenta il tentativo di formulazione della traduzione (*Tentative translation elements* nello schema), un'attività a metà tra l'intuitivo e il riflessivo che mette insieme e rielabora gli input esterni e le informazioni recuperate dalla memoria a lungo termine per produrre un'adeguata traduzione di una data porzione di testo. Alla luce dell'adozione di un CAT Tool nel processo di lavoro, e di quanto appena detto sul recupero delle informazioni dalle risorse digitali, occorre considerare il forte condizionamento esercitato dai suggerimenti forniti dalla TM e dai glossari sulla produzione di tali tentativi di traduzione, fino al punto da rendere in molti casi questo passaggio più un compito del software che del traduttore. Infatti, nelle versioni dei software di qualche anno fa, i segmenti riconosciuti come *exact match* venivano comunque sottoposti al vaglio del traduttore; le versioni più recenti di alcuni CAT Tools, invece, inseriscono automaticamente i match uguali al 100%, senza richiedere alcuna conferma dal traduttore, cosa che se da un lato può avere una grande utilità in fatto di risparmio di tempo e lavoro, può comunque avere i suoi contro perché come ogni traduttore sa, tante volte la stessa frase va tradotta in modo diverso a seconda dei contesti di utilizzo<sup>20</sup>. Allo stesso modo, anche il secondo elemento mediano nel modello di Kiraly ovvero i problemi di traduzione che emergono durante il processo, verranno affrontati tramite il filtro dei suggerimenti del CAT Tool, che proporrà

---

<sup>20</sup> È pur vero tuttavia, che questa opzione può essere disattivata dall'utente se ritenuta più dannosa che benefica.

automaticamente delle possibili soluzioni. Ciò rappresenta sicuramente un'agevolazione per il traduttore dal punto di vista della produttività e della gestione del processo, nonché dal punto di vista della qualità specialmente quando il creatore della TM e dei materiali di riferimento è lui stesso, cosa che gli dà certezza della validità dei suggerimenti ottenuti. Tuttavia, il processo viene ad assumere una certa meccanicità e una forte influenza dalle memorie e i glossari utilizzati, sebbene al traduttore spetti comunque il ruolo decisionale di avallare o rifiutare le proposte ottenute, a tutela della propria libertà professionale. In realtà, come sottolineato da diversi studiosi (Alves e Liparini Campos, 2009) e come si vedrà più avanti, l'influenza prodotta dalle soluzioni proposte porta molto spesso il traduttore a optare per mantenere tali soluzioni senza grandi modifiche, anche laddove non vi sia da parte del cliente o dell'agenzia un'esplicita richiesta di attenersi alle versioni salvate in memoria. Questa forte influenza del contributo della macchina sulla libertà e creatività del traduttore ha una componente sicuramente inconsapevole, non solo laddove si lavori sotto pressione con grossi quantitativi di testi da tradurre in poco tempo e si deleghi per necessità quanto più possibile alla TM. Anche in condizioni di lavoro normali, se la memoria offre un suggerimento il traduttore sarà portato ad adottarlo senza grandi indugi (*ibid.*). Tuttavia, quando la richiesta di attenersi con scrupolo alle scelte linguistiche e alle versioni riportate nella memoria viene dall'esterno, dal cliente o dal project manager, essa può essere accolta in maniera diversa dal traduttore. Già immerso in una comunicazione alquanto insolita nella quale egli non è né il mittente né il reale ricevente ma deve comportarsi contemporaneamente un po' come entrambi (Neubert, in Danks, 1997: 7), costretto non solo a mediare tra due lingue e culture ma anche tra il proprio modo di impostare il lavoro e le direttive ricevute dal PM o semplicemente le esigenze di un lavoro in team, il dover fare i conti con l'ulteriore vincolo rappresentato da TM e terminologia può risultare particolarmente gravoso. Se da un lato il contributo del CAT Tool può far sentire il traduttore maggiormente sicuro nell'utilizzo della terminologia specifica e delle espressioni tipiche dell'ambito linguistico in considerazione, dall'altro può tuttavia capitare (non di rado) che egli non sia d'accordo con le soluzioni adottate nella TM e che gli viene chiesto di utilizzare a sua volta, e che questo causi un senso di frustrazione o di irritazione per il dover limitare la propria libertà e competenza professionale. Ne risulta uno svilimento di quello che è il cuore del modello di Kiraly, il "Translator's self concept", ovvero il modo in cui il traduttore percepisce se

stesso, la propria funzione e il lavoro che sta svolgendo. Alcuni traduttori infatti rifiutano di utilizzare un CAT Tool proprio perché la componente meccanica che essi apportano nel processo traduttivo e il riutilizzo sistematico di quanto già tradotto li porta a percepire il proprio lavoro come un qualcosa di ripetitivo e poco stimolante, in cui le scelte personali e la creatività linguistica vengono mortificate a vantaggio della produttività (Christensen, 2011).

### **3.1.3. Ricadute cognitive dei CAT Tools**

Dalle considerazioni appena compiute utilizzando il modello di Kiraly è possibile comprendere come l'utilizzo dei CAT Tools modifichi il processo traduttivo, l'impostazione del lavoro e la percezione stessa che di esso ha il traduttore. Riproducendo in maniera artificiale alcune delle operazioni mentali del traduttore, essi rendono particolarmente visibile il legame e la reciproca influenza esistente tra processi interni ed esterni, contribuendo così a svelare alcuni aspetti del comportamento cognitivo soggiacente all'attività di traduzione. Nei prossimi paragrafi si farà ricorso a degli studi più particolareggiati per osservare nello specifico alcuni aspetti e passaggi del processo di traduzione particolarmente interessati dall'influenza esercitata dai software di sussidio alla traduzione. Ci si soffermerà nella fattispecie sulla scelta della procedura operativa, sulla scomposizione del testo in unità minime, sull'utilizzo delle memorie e sulla suddivisione dell'attenzione tra le varie parti del processo. Questi aspetti sono stati sviluppati all'interno dei maggiori studi dedicati alle ricadute cognitive degli strumenti di traduzione assistita, ai quali si farà riferimento nei paragrafi seguenti. Tuttavia, prima di affrontare tali studi, è necessaria una premessa relativa agli strumenti utilizzati per poter compiere le osservazioni in essi contenute. Come già detto, comprendere che cosa accade all'interno della mente è alquanto complicato e per molti versi aleatorio, a causa della impossibilità di osservazione diretta di tali dinamiche ma anche della componente personale, emotiva e particolare (legata alle specifiche condizioni ambientali e lavorative) che esercita una grande influenza su di esse. Inoltre, sebbene la ricerca abbia consentito di scoprire molto sul funzionamento del cervello, molti punti rimangono ancora oscuri e di difficile comprensione per gli studiosi delle neuroscienze. Al fine di gettare luce su cosa avvenga nella mente del traduttore durante il suo lavoro, sono stati adottati alcuni metodi di osservazione presi in prestito da altri campi e altre discipline, che osservando alcuni parametri comportamentali consentono di compiere delle supposizioni su cosa

accada all'interno della mente di chi lavora. Tra questi metodi si possono annoverare alcune tecniche computerizzate come il monitoraggio oculare o *eye tracking* (Doherty, O'Brien, Carl, 2010), le riprese video e il monitoraggio dell'attività del computer (utilizzo di tastiera e mouse e screen captures), e alcune forme di verbalizzazione del processo e delle scelte compiute, che possono avvenire in tempo reale come nel caso dei protocolli TAP (*Think-Aloud Protocols*), o successivamente al completamento della traduzione. Ognuna di queste tecniche serve a far luce su aspetti diversi, e ognuna di esse possiede vantaggi e svantaggi, per cui non di rado esse vengono utilizzate in maniera congiunta in modo da avere un riscontro dei risultati prodotti. Nei prossimi paragrafi si descriverà più nel dettaglio il loro utilizzo all'interno di studi empirici ed il tipo di informazioni che essi possono fornire.

### 3.2 *Definizione di una procedura operativa*

Una delle prerogative che distinguono i traduttori professionisti da quelli alle prime armi è l'aver sviluppato negli anni di esperienza un iter operativo che regolamenti il loro lavoro e al quale attenersi per avere un perfetto controllo delle varie attività da svolgere. In generale, tale iter o procedura prevede operazioni da svolgere prima di iniziare la traduzione (*pre-translation*), la preparazione dei materiali da consultare, una serie di regole per gestire la traduzione propriamente detta, la rilettura e revisione del testo tradotto ed infine una serie di operazioni di riordino e gestione dei documenti utilizzati e prodotti (*post-translation*). Sebbene tali attività siano grosso modo standardizzate, ogni traduttore organizza il proprio lavoro in base a criteri e preferenze personali. Laddove si utilizzi un CAT Tool, questo entrerà a far parte della procedura operativa determinando diverse delle operazioni al suo interno. Infatti, come già precedentemente accennato, alcuni traduttori che hanno lavorato per anni senza alcun sussidio per la traduzione, rifiutano o trovano molto scomodo iniziare ad utilizzare un CAT Tool sia per la necessità di doverne imparare l'uso, sia per il fastidio di dover modificare una procedura operativa già collaudata da anni. I nuovi traduttori che invece sono entrati in contatto con i CAT Tools già durante la loro formazione accademica, sono più inclini ad adottarli nella propria attività, e più pronti all'evenienza di dover imparare l'utilizzo di specifici strumenti in base alle richieste delle agenzie con le quali si troveranno a collaborare. In effetti, come si vedrà nella terza parte della presente tesi,

L'affermazione degli strumenti di traduzione assistita negli ultimi vent'anni ha cambiato o meglio ampliato l'ambito delle competenze professionali necessarie per tradurre, con l'introduzione nella didattica di un'attenzione allo sviluppo delle competenze tecniche - informatiche (Pym, 2003: 483).

### **3.2.1. I CAT Tools strumenti di organizzazione del lavoro**

L'uso di un CAT Tool segue il lavoro di traduzione dall'inizio alla fine, dando vita a un processo fatto di tappe precise che si susseguono in maniera ordinata o si richiamano in un continuum atto a rendere il massimo dal punto di vista qualitativo e quantitativo. Come emerge dagli studi di numerosi ricercatori (Dragsted 2004 e 2006; Christensen 2011; Christensen e Schjoldager 2010; O'Brien 2006, 2008 e 2009; O'Brien, O'Hagan e Flanagan 2010; Alves e Liparini Campos 2009 etc.) le modifiche al processo traduttivo causate dall'applicazione di tali strumenti sono conseguenza di una serie di operazioni cognitive prima che pratiche. Come puntualizzato da Disparti (2004: 52-53),

*“L'utilizzo del computer si accompagna a variazioni cognitive ed operative negli utenti: non si tratta solo di dati ma anche di abilità metacognitive che trovano espressione nelle attività mentali implicate. Ogni fase si differenzia dalle altre per le modalità operative e per il diverso equilibrio e le differenti modalità procedurali degli schemi cognitivi. (...) Le modifiche sostanziali riguardano il passaggio da un'intelligenza quasi esclusivamente linguistica e logica ad una integrata, che equilibra le precedenti con quella visiva (...). Quasi un uso consapevole di puzzle di abilità cognitive e di aree di conoscenze, fatto di utilizzi immediati e di rinvii”.*

Quanto affermato da Disparti sull'adozione del computer in maniera più generica nelle attività quotidiane si può ricondurre al caso più specifico in questione, l'introduzione dei CAT Tools in un processo traduttivo professionale. I cambiamenti sono della stessa natura, cognitivi prima che operativi, e riguardanti il modo in cui il lavoro viene percepito (una prospettiva psicologica e semantica oltre che cognitiva), l'organizzazione mentale del lavoro stesso con la messa in gioco di abilità diverse. È anche una questione metacognitiva, laddove il traduttore percepisce una differenza nel

proprio modo di operare o nella sua capacità di gestire ed utilizzare le proprie conoscenze e competenze, e sa valutare la maniera migliore di mettere a frutto i propri processi cognitivi per massimizzare il risultato che intende perseguire. I software di sussidio alla traduzione semplificano e accelerano il processo, ma rendono più complesso d'altro canto l'insieme delle competenze e delle intelligenze che evocano, non solo linguistiche ed operative, ma anche multimediali, organizzative e di sintesi, per mettere insieme i vari stimoli ottenuti dai diversi strumenti utilizzati. La mente umana percepisce questa maggiore complessità di operazioni e le differenze che essa comporta nello svolgimento del compito, tuttavia è abile nel creare armonia tra le sue varie parti e tra le conoscenze che possiede, al fine di produrre soluzioni e risolvere problemi. Tale processo è per metà consapevole e per metà inconscio, come una risposta immediata a una situazione sulla quale però l'individuo è in grado di riflettere e ragionare (Shreve, in Danks *et al.* 1997).

### **3.2.2. Collaborazione uomo - computer**

Se è vero quanto sopra affermato, allora la semplice introduzione di una componente computerizzata con la quale interfacciarsi implica già un diverso approccio e dei processi diversi. Affidando parte del proprio lavoro alla macchina e lasciandosi guidare dal programma tappa per tappa, il traduttore entra in una dinamica di cooperazione uomo – computer che può essere considerata come un insieme di abilità ed operazioni cognitive distribuite (*distributed* o *collaborative cognition*, Harnard 2005, cit. in Dragsted 2006) tra la mente umana e l'artefatto digitale o tra più individui chiamati a interfacciarsi per tramite della TM che condividono. Secondo Hutchins (1995), i processi cognitivi nascono e si sviluppano in stretto legame con l'ambiente in cui l'individuo agisce e con le attività che egli svolge in esso, ed in maniera condivisa e collaborativa con tutti gli altri individui operanti in tale ambiente. Le rappresentazioni mentali di tali contesti operativi, prodotte dai processi cognitivi messi in atto, non possono quindi che avere una natura condivisa e distribuita tra i membri dello stesso gruppo sociale. In questa prospettiva, gli strumenti computerizzati modificano l'ambiente originario introducendo una componente tecnologica, ma nello stesso tempo divengono strumento di creazione e condivisione di nuove rappresentazioni (Hutchins, 2001). Secondo Zhang (cit. in Dragsted, 2006), tali rappresentazioni delle realtà esterne sono talmente forti e radicate nella mente al punto da influenzare e persino determinare il comportamento cognitivo. Un CAT Tool indubbiamente modifica già a partire dall'aspetto visuale, ovvero al livello

di rappresentazione esterna, l'approccio al processo di traduzione, includendolo in un'interfaccia che lo presenterà secondo determinati criteri di presentazione delle risorse e scomposizione del testo che potrebbero non necessariamente corrispondere con la rappresentazione mentale propria del traduttore. Ciò ha sicuramente un impatto su alcuni aspetti comportamentali relativi al processo traduttivo e sulla percezione che di esso ha il traduttore. Per quanto personalizzabile possa essere la struttura del software, e per quanta libertà esso possa lasciare all'utente di avvalersi o no di alcune risorse e opzioni, si verrà comunque a creare una sovrapposizione di modalità operative che dovranno trovare una sintesi nella creazione, da parte del traduttore, di una nuova rappresentazione di sé e del proprio lavoro, non solo da un punto di vista operativo ma anche da quello percettivo, che tenga conto dello strumento adottato (Wallis, 2006).

Da un punto di vista più pratico, sono diversi gli aspetti della nuova modalità di lavoro che possono essere interessanti dal punto di vista cognitivo. Per esempio, la ricerca delle informazioni necessarie all'interno di un unico ambiente di lavoro piuttosto che in diverse finestre, pagine ed applicazioni, ha il vantaggio di una minore dispersione dell'attenzione e di una maggiore armonizzazione tra le varie risorse adoperate, che possono venire confrontate con facilità, favorendo ed accelerando i processi di *problem-solving*. Laddove vengano utilizzati all'interno del CAT Tool due glossari, uno proprio e uno estratto da internet, i risultati verranno visualizzati all'interno della stessa cartella ma differenziati in modo da distinguerne l'origine. Il loro confronto, unito magari alla ricerca tramite *Concordancer* o traduzione automatica, sempre visualizzabili all'interno dell'unica schermata, consentirà al traduttore di avere una panoramica di suggerimenti davanti ai suoi occhi, tra i quali poter scegliere in base al contesto o alle esigenze del documento. Un altro aspetto che dimostra di avere un'influenza sulla percezione del lavoro è la possibilità di avere una visuale chiara del carico di lavoro da svolgere. L'analisi compiuta dal programma rivela quanto già di esso è stato tradotto o è rispondente alle risorse presenti in memoria. Ciò predispone il traduttore in maniera più consapevole nei confronti del lavoro che dovrà affrontare, lo aiuterà a gestire i suoi tempi in base alla data di consegna e gli permetterà di monitorare costantemente i propri progressi.

### 3.2.3. Impatto sui tempi di lavoro

L'adozione di un CAT Tool comporta, come dimostrato da alcune osservazioni ed esperimenti condotti da Dragsted (2004), Christensen e Schjoldager (2011) ed altri, una diversità nelle tempistiche di lavoro, sia totali che parziali, ovvero relative alle diverse fasi del lavoro. Dragsted (2004) riporta i risultati di un esperimento condotto con 6 traduttori professionisti e 6 studenti, ai quali è stato chiesto di tradurre due testi, simili in contenuto e lunghezza. Il primo doveva essere tradotto nella maniera classica, il secondo utilizzando un CAT Tool. Tra il primo e il secondo test è stato fatto passare del tempo, per evitare che il recente uso della terminologia comune influenzasse lo svolgimento della seconda fase dell'esperimento. I risultati mostrano in generale una riduzione dei tempi di produzione nella seconda fase, quella in cui è stato utilizzato il CAT Tool. Tuttavia, mentre fra i traduttori professionisti la riduzione si è dimostrata minima e in alcuni casi i tempi si sono persino prolungati, con una media finale pari al -2%, tutti gli studenti hanno prodotto tempi significativamente più brevi, con una media del -16%. Tale differenza rispecchia il diverso grado di affidamento fatto dai partecipanti nei confronti degli strumenti utilizzati, come si deduce anche dal questionario di valutazione nel quale gli studenti hanno affermato nella maggior parte dei casi di aver accettato i suggerimenti della memoria senza troppo soffermarsi su di essi, mentre dalle risposte dei traduttori professionisti si riscontra un maggiore bisogno di "controllo" del processo, che li portava a verificare e valutare i suggerimenti, apportando se del caso delle modifiche.

Anche i tempi dedicati alla revisione si sono dimostrati suscettibili di modifiche con l'introduzione dei CAT Tools, in maniera persino più significativa dei tempi di traduzione. Da un successivo esperimento di Dragsted (2006), condotto secondo lo stesso criterio di quello sopra descritto, si è potuto constatare che i tempi di revisione si riducevano in media del 34% per i traduttori professionisti, e persino dell'84% per gli studenti. Tramite l'utilizzo di *Translog*, un software che consente di registrare e analizzare l'attività di tastiera e mouse svolta durante il processo di traduzione e i tempi di pausa tra di essi, si è riscontrato che anche senza l'uso del CAT Tool i traduttori professionisti dedicassero a ogni segmento minor tempo rispetto agli studenti durante la fase di traduzione, non soffermandosi per esempio a considerare i problemi lessicali e lasciando quindi la soluzione di eventuali difficoltà alla fase di revisione. Gli studenti, dal canto loro, hanno invece mostrato la tendenza di soffermarsi maggiormente sui segmenti

in fase di traduzione, cercando di risolvere i problemi man mano che li incontravano. Occorre quindi considerare già a monte una differenza dei tempi di revisione – più lunghi per i professionisti, più snelli per gli studenti – che l'introduzione del CAT Tool ha contribuito in entrambi i casi a ridurre in proporzioni diverse che rispecchiano il comportamento e le tendenze naturali degli intervistati.

#### **3.2.4. Le impressioni degli utenti**

Oltre all'adozione di software specifici come *Translog*, il questionario di valutazione è una delle tecniche di osservazione utilizzate negli studi che mirano a comprendere i processi cognitivi. Esso fa parte dei metodi di enunciazione dei processi, che possono prevedere una narrazione vocale o per iscritto, in tempo reale o differito, del compito svolto, con un'attenzione agli aspetti particolari, ai motivi che hanno spinto verso le scelte adottate, alle difficoltà incontrate, e con una valutazione soggettiva degli aspetti positivi o negativi, vantaggi e svantaggi ecc. Dal questionario dell'esperimento di Dragsted già citato, è possibile ricavare delle informazioni anche sulle capacità di concentrazione dei partecipanti. Diversi infatti sostengono che il dover lavorare in maniera ordinata segmento per segmento li ha aiutati a concentrarsi sulle frasi procedendo in maniera ordinata senza disperdere l'attenzione in altre parti del testo. Tuttavia, la divisione in segmenti è allo stesso tempo vista in maniera negativa da alcuni partecipanti, per via dell'eccessiva rigidità che comporta e della mancanza di una visuale d'insieme del testo. Si lamenta in particolare il rischio di perdere informazioni contenute nel testo e fondamentali sin dall'inizio della traduzione, e quello di non riuscire a rendere armonico il proprio lavoro per via della divisione in frasi del testo originario, difficilmente riproducibile nel testo di arrivo. Si pensi alla differenza tra la formulazione dei periodi in inglese e in italiano. La prima lingua predilige frasi più brevi e poco complesse, mentre nella seconda sono frequenti periodi più complessi composti da più frasi subordinate o coordinate. La necessità di riformulare concetti espressi in più frasi può rendere la segmentazione frustrante per il traduttore e inficiare la qualità finale del lavoro. In ogni caso, anche laddove questa difficoltà non ci fosse, l'ordine sistematico dato dalla segmentazione, il dover procedere secondo il rigido schema del software e la ridotta libertà nell'approccio col testo richiedono, a fronte dei benefici che ne possono derivare, una capacità e una volontà di adattamento del proprio personale approccio al lavoro e all'esplorazione del testo.

Infine, un aspetto negativo evidenziato da diversi utenti dei CAT Tools è la complessa gestione dei molteplici file che il programma crea per salvare le risorse adoperate e prodotte. Non si tratta solo dei documenti di partenza e di arrivo, ma anche di file di sistema, file log e tutta una serie di file autoprodotti, la cui perdita o danneggiamento accidentale può compromettere l'utilizzo dei materiali e delle risorse. Ciò richiede un ordine meticoloso con la creazione e manutenzione di cartelle e sottocartelle specifiche, che seguano un criterio di catalogazione ben preciso con cartelle separate e diciture specifiche per ogni progetto. Se questo può apparire da un lato un elemento fondamentalmente pratico e poco rilevante, dall'altro è possibile ricollegarlo a una competenza più profonda, ovvero la capacità gestionale e organizzativa del traduttore, che ricalca in maniera visibile e concreta l'ordine mentale che egli si crea, le sue mappe operative mentali, la propria rappresentazione delle conoscenze e la capacità di gestire i processi e le risorse mentalmente prima che praticamente (Dancette e Halimi, 2005).

### 3.3 *La scomposizione del testo*

Si è detto che i CAT Tools, nello specifico le memorie di traduzione, suddividono i testi in porzioni minime che vengono poi accoppiate con la rispettiva traduzione per formare delle *Translation Units* (TU). Nella schermata di lavoro, il traduttore si troverà quindi ad operare su di un testo segmentato secondo il criterio adoperato dal software, e sottopostogli in maniera tale che egli possa lavorare esclusivamente su un segmento alla volta. In genere, i CAT Tools danno la possibilità di impostare il livello testuale al quale segmentare il testo, scegliendo tra la fine di ogni frase o paragrafo o in base alla punteggiatura. Tuttavia, solitamente l'impostazione di default è la divisione in frasi, riconosciute per comodità come l'unità sufficientemente lunga da avere un senso compiuto e al contempo sufficientemente breve da poter essere elaborata dalla memoria a breve termine come un tutt'uno concettuale e formale (Alves e Couto Vale, 2009).

#### **3.3.1. La segmentazione frase per frase**

Come si è accennato nel paragrafo precedente, malgrado il principio che ha stabilito l'affermazione della frase come unità minima o *Translation Unit*, e malgrado la sua adozione da parte di tutti i CAT Tools e la maggioranza degli utenti, numerosi

ricercatori si sono interrogati sulla sua reale corrispondenza con il livello di scomposizione mentale dei testi (la “segmentazione cognitiva”, Dragsted 2004: 6). Già a partire dai primi decenni del secolo scorso la definizione di TU divenne una delle questioni centrali nei Translation Studies, nel tentativo di identificare quale fosse, all’interno del testo, l’unità minima di senso che venisse adottata dal traduttore per trasporre il contenuto del documento nella lingua di arrivo. Essa fu inizialmente identificata con la parola, ovvero con l’unità lessicale minima, di cui era possibile trovare un equivalente grammaticale e di significato. La frase, dal canto suo, fu meramente vista come una combinazione logica di parole fino agli anni ’50, quando le teorie della Scuola di Praga e la grammatica generativa di Chomsky (1957) enfatizzarono l’importanza della frase come costituente minima di senso, sulla base del fatto che i traduttori non traducono “parole”, ma “unità di pensiero” (Vinay and Darbelnet 1995; cit. in Dragsted 2004) espresse attraverso costrutti articolati, la cui forma minima era appunto la frase. Un’ulteriore prospettiva venne aperta dagli studi sul *discourse analysis*, che considerarono il testo nella sua interezza come unica unità di senso (Bassnett, 1991), alla cui formulazione contribuiscono non solo il contenuto ma anche lo stile e la struttura complessiva. Il contributo di tutti questi approcci condusse all’affermazione dell’idea che il concetto di unità di traduzione fosse tutt’altro che univoco e chiaramente definibile, ma al contrario un elemento dalla grandezza variabile, definito come “*the minimal stretch of language that has to be translated together as one unit*” (Newmark 1988: 54); porzione che deve essere “più piccola possibile e sufficientemente grande” (Newmark 1988:285) a stabilire un’equivalenza traduttiva e che può quindi essere composta da una parola, una collocazione, una frase o un intero paragrafo. Secondo Ballard (2010), un’unità di traduzione viene generata dall’applicazione di una strategia traduttiva su una porzione di testo, e l’interruzione tra un’unità ed un’altra viene proprio determinata dal cambio di strategia adottata. Tale approccio esclude quindi la divisione “quantitativa”, definendo un approccio diverso, più funzionale, alla questione. Interessante è infine notare il punto di vista della scuola cinese, rappresentata da Luo (1999, cit. in Dragsted, 2004). Egli afferma che prima di parlare di unità di traduzione occorre distinguere tra *unità di analisi* ed *unità di trasposizione*, ovvero tra la porzione minima a comprendere il senso e il contenuto del testo, e la porzione di testo e di senso che renda più facile l’identificazione di un equivalente nella lingua di arrivo. Tali unità non sono commensurabili, occorrerà un’unità

di testo più grande per l'analisi, mentre sarà necessario soffermarsi su segmenti più piccoli per poter operare una traduzione accurata.

### **3.3.2. L'unità cognitiva di traduzione**

Ai fini della presente tesi, tuttavia, oltre e più della definizione linguistica di TU ci interesserà comprendere quella prodotta adottando una prospettiva cognitiva. Fu Sager (1994) per primo a realizzare che l'unità di traduzione avesse una duplice natura, linguistica da un lato e cognitiva dall'altro, e a collegare quest'ultima alla capacità della mente umana di elaborare insieme una quantità limitata di informazioni. Secondo Sager, quindi, il concetto di TU assume anche un connotato individuale e soggettivo, dipendente dall'abilità della memoria a breve termine di gestire quantità più o meno ampie di informazioni, e dai criteri individuali spontanei o "pensati" che determinano le modalità di suddivisione del testo. Egli inoltre considera anche il grado di familiarità del traduttore con l'argomento e la tipologia testuale, osservando come le unità di traduzione siano più grandi se il traduttore è specialista nel dato linguaggio o argomento, mentre tenderanno a ridursi nel caso laddove i traduttori siano meno esperti.

Il tentativo di comprendere il comportamento cognitivo dei traduttori ha reso necessario lo sviluppo di appositi studi empirici attraverso i quali provare a desumere, tramite esperimenti e osservazioni, la corrispondenza delle teorie linguistiche con i processi mentali messi in atto dai traduttori. Al fine di "aprire un varco" nel difficilmente osservabile processo cognitivo, sono state adottate delle modalità di indagine che registrano alcuni parametri comportamentali in base ai quali giungere a delle conclusioni sui meccanismi interni correlati. Tra questi, si è già fatto riferimento al software Translog, per la registrazione dell'attività di tastiera e mouse e della durata delle pause. Altri strumenti vengono utilizzati anche per monitorare il movimento degli occhi (eye-tracking) o, in alcuni casi, per registrare il comportamento neurale. Oltre agli strumenti tecnologici, il metodo principale adoperato per raccogliere informazioni è rappresentato da alcune forme di verbalizzazione quali i *Think-Aloud Protocols* (Bernardini, 1999 e 2001) ed altre modalità di descrizione orale o scritta delle procedure seguite e delle scelte compiute. Tale verbalizzazione può avvenire simultaneamente al lavoro (*online data*, cit. Christensen 2011:3), o in maniera retrospettiva, successivamente ad esso (*offline data*, *ibid.*). Ericsson e Simon, che per primi hanno descritto e sperimentato i protocolli TAP, ne hanno anche delineato i limiti, principalmente identificabili col fatto che essi danno

accesso solo alle informazioni elaborate dalla memoria a breve termine e trattenute per un tempo sufficiente ad essere espresse verbalmente, mentre non rendono possibile l'accesso alle attività cognitive più profonde e talmente radicate da essere compiute in maniera automatica (Ericsson e Simon, 1984: 15). Inoltre, la validità e affidabilità dei dati raccolti è stata anche messa in discussione da diversi ricercatori, che sostengono che i processi soggiacenti alla determinazione di scelte, giudizi e comportamenti non possono essere scoperti tramite introspezione diretta o comunque non compresi in maniera totale solo tramite verbalizzazione (Misbett e Wilson, Seliger, Börsch cit. in Dragsted, 2004). Malgrado ciò, i TAP sono stati adoperati in diverse sperimentazioni, e laddove possibile o necessario, sono stati affiancati da altre metodologie e strumenti, al fine di confrontare i dati raccolti e potere quindi giungere a conclusioni sufficientemente comprovate.

Diversi studi sono stati condotti a partire dagli anni '80 al fine di osservare quale fosse l'unità di traduzione "naturale". La prassi comune di tali studi prevedeva lo svolgimento di una traduzione da parte di un gruppo di partecipanti, nella maggior parte dei casi studenti di traduzione, e la successiva o simultanea verbalizzazione introspettiva circa il metodo seguito e le azioni svolte passo dopo passo. I risultati hanno confermato che il concetto di TU è qualcosa di molto ampio e variabile, solitamente una frase o enunciato di 2-6 parole circa (Gerloff 1986; Lörcher 1986; Krings 1986b; Kiraly 1995, cit. in Dragsted 2004: 147), ma comunque dipendente da alcune variabili individuali. Inoltre, è stato ipotizzato che le TU possano essere costruite secondo una scala gerarchica, con un'unità principale solitamente corrispondente alla frase, e delle sotto-unità al suo interno (Bernardini 1999: 185), in linea con quanto suggerito da Luo (1999, cit. in Dragsted 2004), già sopra citato.

Gli studi a cui si fa riferimento hanno anche utilizzato la registrazione delle pause e delle esitazioni, che pur essendo principalmente riconosciute come identificatrici di difficoltà o incertezze nello svolgimento della traduzione, hanno anche dimostrato di poter fornire informazioni anche sull'organizzazione mentale che il traduttore sviluppa relativamente al testo originario. Delle pause sufficientemente lunghe possono essere riconosciute come possibili indicatori del punto in cui termina una TU e ne comincia un'altra.

Con l'introduzione delle TM e dei CAT Tools, basati sulla scomposizione dei testi in segmenti, la discussione sull'identificazione della "TU" vista come unità cognitiva si è arricchita di questo nuovo elemento, che ha indirizzato gli studi verso la valutazione

delle possibili ripercussioni dell'adozione dei sistemi CAT e della loro rigida segmentazione in frasi sul processo traduttivo e sul prodotto finale. Negli ultimi 8 anni circa sono state condotte delle sperimentazioni in tal senso, che hanno mostrato in tutti i casi una difficile corrispondenza dei livelli di analisi ed elaborazione dei testi dei partecipanti con lo schema operativo dei CAT Tools, basato sulla frase. Christensen (2011) riporta l'esito di uno studio nel quale 23 studenti hanno prima seguito un corso teorico – pratico di traduzione assistita da computer durante il quale hanno dovuto completare delle traduzioni utilizzando una TM, e appena completato il corso hanno compilato un questionario per raccogliere le loro impressioni a riguardo. Al di là delle valutazioni positive sui benefici derivanti dall'uso del CAT Tool, proprio relativamente al rigido focus sulla frase come unità operativa, più della metà degli intervenuti ha espresso un parere negativo a riguardo. Pur riconoscendo che la divisione in frasi aiuta la concentrazione e la gestione del lavoro, è opinione comune che essa possa rendere più difficile la comprensione del contesto generale, e rappresenti praticamente o anche solo percettivamente un vincolo per coloro che sono abituati a lavorare con porzioni di testo più piccole o più grandi della frase.

### **3.3.3. Studi sull'argomento**

Lo studio di Dragsted (2004) al quale si è già fatto riferimento ha coinvolto sia studenti di traduzione che traduttori esperti. Esso si è composto di due parti, una prima dedicata all'osservazione della traduzione umana attraverso la raccolta dei dati con Translog; la seconda parte incentrata sulla sperimentazione dell'uso dei CAT Tools in traduzione tramite verbalizzazione retrospettiva. L'intenzione di Dragsted era quella di comprendere se la segmentazione in frasi corrispondesse all'unità cognitiva di segmentazione, e in un secondo momento ipotizzare una segmentazione alternativa e più funzionale. Dai risultati ottenuti, si è dedotto che la frase non corrisponde all'unità principale della segmentazione cognitiva, specialmente per i traduttori professionisti, che mostrano più nettamente l'esigenza di muoversi più liberamente nel testo, prendendo in esame ora porzioni più piccole, ora parti più estese di testo. Dalle loro risposte al questionario essi appaiono maggiormente consapevoli rispetto agli studenti di come la segmentazione frase per frase adottata dal CAT Tool abbia un effetto vincolante sul loro comportamento cognitivo e sulle rappresentazioni mentali del testo, modificando in ultima istanza il lavoro di traduzione (Dragsted 2004: 129). Anche la registrazione delle

pause all'interno di un unico segmento suggerisce che sia i professionisti che gli studenti abbiano bisogno di concentrarsi su porzioni più piccole di testo, mentre l'attività del mouse registrata con Translog rivela un andare avanti e indietro nel testo per verificare quanto già tradotto o ricercare nel testo, soprattutto quello immediatamente più prossimo, informazioni necessarie a una corretta traduzione. Ciò sembra confermare le conclusioni di Dragsted sulla non corrispondenza della frase con l'unità di segmentazione cognitiva, conclusioni che spingono l'autrice a proporre l'adozione di una segmentazione al di sotto del livello della frase, che oltre a fornire verosimilmente maggiori risultati dalla consultazione della TM, corrisponderebbe maggiormente ai processi naturali di segmentazione compiuti dalla mente del traduttore (Dragsted 2004: 240).

Nel suo studio del 2008, O'Brien adotta un punto di vista diverso ma utile comunque a gettare luce sul rapporto tra segmentazione della macchina e segmentazione cognitiva. L'autrice infatti tenta di comprendere quale sia il grado di apprezzamento degli utenti dei CAT Tools circa la funzione "Concordance", che serve a cercare porzioni di testo contenute all'interno di un segmento, ovvero di una frase. Il *Concordance search* è una delle funzionalità maggiormente usate dai traduttori, a conferma dell'attitudine di identificare all'interno dell'unità di traduzione (o TU) delle sotto-unità costituite da parole o raggruppamenti di parole (i *sub-segment matches* come vengono definiti in Bowker e Fisher, 2010: 62). La funzione *Concordance* può essere mantenuta attiva sempre durante la traduzione, in modo che laddove non si rinvenga un match essa possa automaticamente fornire delle informazioni utilizzabili dal traduttore ricercando le singole parole costituenti la frase e visualizzando per ciascuna i vari utilizzi rinvenuti nella memoria. In alternativa, essa può essere utilizzata laddove se ne avverta l'utilità per ricercare determinate espressioni all'interno del testo. Nell'esperimento di O'Brien un gruppo di partecipanti ha tradotto con la finestra del Concordancer sempre attiva, mentre un altro gruppo senza di esso. Attraverso l'utilizzo di un dispositivo di monitoraggio oculare, si è osservata la frequenza e la durata della fissazione sulla finestra del Concordancer, mentre con uno strumento di controllo dell'attività di mouse e tastiera si è valutata la frequenza delle operazioni di copia e incolla da questa finestra. Le conclusioni mostrano un massiccio utilizzo della funzionalità attivata, che implica anche un accrescimento dei tempi totali di traduzione ma d'altro canto anche un notevole miglioramento della qualità finale della traduzione (monitorata con un software di Quality

Assessment). Ciò suggerisce che i traduttori sono consapevoli dell'importanza di concentrarsi sulle sotto-unità all'interno della TU, e del bisogno di valutare attentamente l'utilizzo della porzione di testo ottenuta dal Concordancer all'interno della frase interessata, come dimostra l'accrescimento dei tempi di lavoro riscontrato dall'esperimento di O'Brien (2008).

Un'ultima considerazione viene invece da Mossop (2006), il quale invita a valutare come la segmentazione del testo implichi il rischio di perdita di una visione d'insieme del testo come unità semantica e l'insorgenza di un fenomeno che lui definisce "traduzione collage", con riferimento anche all'abitudine di "smembrare" un testo in parti da assegnare a traduttori diversi, che determina la necessità di "ricucire insieme" i vari pezzi una volta completati. Mossop sostiene la necessità di una maggiore armonizzazione delle parti e la valorizzazione del testo nella sua interezza, contro la diffusa tendenza di suddividere il testo, giustificata prevalentemente da motivi economici e commerciali.

### 3.4 *La memoria, le memorie*

Dedichiamo adesso un paragrafo all'osservazione del funzionamento della memoria umana, o per meglio dire delle *memorie* che regolano i diversi processi mentali e cognitivi, per poter tracciare in un secondo momento analogie e differenze con le memorie di traduzione digitali, e tentare di determinare se l'introduzione di una TM nel processo traduttivo comporti delle modifiche alle capacità e alle prestazioni della memoria, in particolare relativamente all'acquisizione di nuovi dati e al recupero delle informazioni già possedute. Questo aspetto, a dire il vero, non è stato sinora al centro di molti studi e ricerche, quindi si attingerà dagli studi relativi all'organizzazione della memoria umana e si cercherà di trarre da tali basi delle considerazioni utili ai nostri scopi. Prima di addentrarci nell'esplorazione dei processi mnesici relativi alla traduzione, occorrerà definire i concetti fondamentali come la suddivisione dei vari tipi di memorie ed i compiti che ciascuna di esse svolge. Come suggerisce Dragsted (2004: 33), il processo traduttivo può essere assimilato al processo umano di elaborazione delle informazioni, quindi ad esso si applicheranno le stesse teorie sui processi mnesici valide per questo più ampio ambito di competenza.

### **3.4.1. Le diverse *memorie***

Come è ben noto, le informazioni elaborate dal cervello umano vengono immagazzinate da diverse memorie, principalmente riconducibili a tre grandi categorie: la memoria sensoriale, la memoria a breve termine o di lavoro (anche se come si vedrà questi termini non esprimono esattamente lo stesso concetto) e infine la memoria a lungo termine. Tali memorie sono diverse non solo per i compiti che svolgono, ma anche per la durata di ritenzione dell'informazione e per la quantità di informazioni che sono capaci di gestire.

La memoria sensoriale riceve le informazioni dagli organi percettivi e le trattiene appena per delle frazioni di secondo. Essa ha inoltre una capacità di ricezione molto selettiva per via delle molteplici sollecitazioni sensoriali che riceviamo dall'esterno, e che devono necessariamente essere schermate perché il cervello riesca a svolgere i suoi compiti senza lasciarsi distrarre dalle interferenze sensoriali provenienti dall'ambiente circostante. Tale schermatura è ancora maggiore nel caso in cui la persona sia impegnata a svolgere un'attività mentale complessa, che richiede tutta la sua concentrazione. Una volta che la memoria sensoriale ha selezionato le informazioni, le passa alla memoria a breve termine, definita classicamente come un deposito immediatamente accessibile destinato all'immagazzinamento temporaneo delle informazioni che devono essere elaborate in maniera consapevole al fine di eseguire una qualsiasi attività cognitiva (Baddeley 1992: 34). La memoria a breve termine (MBT) riceve input dalla memoria sensoriale e da quella a lungo termine, e li utilizza per produrre un output (Dragsted 2004: 36), per questo è stata anche definita come "the workspace of the mind", la stazione di lavoro della mente. Il termine "memoria di lavoro" (MdL) è stato adottato per la prima volta al posto di "memoria a breve termine" da Baddeley e Hitch (1974) non tanto come mero sostituto della precedente nomenclatura, quanto con l'intenzione di introdurre un modello della memoria a componenti multiple, in contrasto col modello precedente che considerava la memoria a breve termine come un unicum. Baddeley (1992) definisce la memoria di lavoro come un sistema dalle capacità limitate che permette l'immagazzinamento momentaneo e l'elaborazione delle informazioni necessarie alla realizzazione di compiti cognitivi complessi, come la comprensione, l'apprendimento e il ragionamento. Secondo il modello di Baddeley e Hitch (1974), la memoria di lavoro si compone di tre parti principali, un sistema di controllo centrale che supervisiona e

coordina le funzioni degli altri due sotto-sistemi ausiliari, ovvero il *loop fonologico*, responsabile dell'immagazzinamento e dell'utilizzo delle informazioni provenienti dal linguaggio (formazioni verbali o uditive), e il "*taccuino*" *visuo-spaziale* considerato responsabile del mantenimento e dell'elaborazione di informazioni visuo-spaziali, nonché della capacità di generare immagini mentali. A questi due sotto-sistemi è stato aggiunto sempre da Baddeley nel 2000 un terzo elemento, il *buffer episodico*, che rappresenta un sistema di memoria separato, di capacità limitata, e che utilizza una codifica di tipo multimodale, per creare collegamenti tra informazioni che sono integrate da sottosistemi con codici diversi, dando vita a rappresentazioni unitarie significative e coerenti, ovvero scene ed episodi. Questo sistema a tre sotto-livelli coordinati dall'unità centrale sarebbe implicato, secondo Baddeley, nella realizzazione dei compiti cognitivi più complessi, come la traduzione nel nostro caso, e di tutte le attività mentali consapevoli implicanti l'acquisizione di nuove conoscenze, la definizione di decisioni e soluzioni a problemi.

Una caratteristica molto importante della memoria di lavoro è la ridotta capacità nel numero di elementi che essa riesce a trattenere, nonché la limitata durata di ritenzione delle informazioni, pari a circa 20 secondi o persino meno. L'influenza prodotta dalle interferenze che essa intercetta nel frattempo, come ad esempio l'aggiungersi di altre informazioni o semplicemente il passare del tempo, causa il decadimento dei dati precedentemente focalizzati. Per evitare di dimenticare informazioni importanti, occorrerà tornarvi spesso in maniera reiterata perché esse si fissino nella memoria a lungo termine. L'operazione di ripetizione può avvenire volontariamente, proprio al fine di "imparare a memoria" qualcosa, o anche involontariamente, in maniera automatica senza che la persona ne sia cosciente, per esempio quando si elabora la stessa informazione per diverse volte fino a quando essa viene interiorizzata definitivamente.

Entra così in gioco la memoria a lungo termine (MLT), il magazzino potenzialmente illimitato sia per la quantità di informazioni che può contenere che per la durata della conservazione. Si tratta di una memoria permanente (ovvero che può trattenere i dati in maniera definitiva) che però, al contrario della MBT e della memoria sensoriale, comporta un accesso più lento alle informazioni, che nella maggior parte dei casi hanno bisogno di stimoli esterni o interni per essere evocate. Inoltre, l'accesso a tali informazioni può avvenire comunque solo attraverso la memoria di lavoro, unico livello,

come si è già detto, a fornire consapevolezza dell'informazione posseduta. La memoria a lungo termine viene divisa in memoria episodica e memoria semantica. La prima è il bagaglio delle nostre esperienze ed eventi vissuti, in altre parole in nostro passato. La seconda invece è un magazzino strutturato che ospita concetti, abilità e conoscenze che si sono acquisite. In qualche modo la memoria semantica dipende da quella episodica, dato che le nostre conoscenze sono frutto delle esperienze vissute.

### **3.4.2. Memoria e traduzione**

Senza dilungarsi oltre sulla descrizione delle memorie, si passerà adesso a vedere come esse determinino il processo traduttivo. Esso è frutto dell'interazione tra la memoria di lavoro e quella a lungo termine, il risultato delle conoscenze possedute e delle capacità di elaborazione e produzione della mente. Più delle altre, ci interesserà la memoria di lavoro o operativa, essendo lei la responsabile dell'aspetto dinamico del processo, composto dall'acquisizione e gestione delle informazioni e dai processi decisionali e risolutivi.

La traduzione come attività mentale comporta diverse funzioni cognitive, le principali delle quali sono la lettura, la comprensione e infine la redazione. Riprendendo la divisione in livelli di Baddeley e Hitch (1974), possiamo dire che il loop fonologico ed il taccuino visuo-spaziale sono responsabili dell'elaborazione delle informazioni visive e verbali recepite con la lettura, mentre la comprensione necessita dell'interazione tra la memoria di lavoro e la memoria a lungo termine e la formazione di immagini mentali legata agli stessi sottolivelli ma anche al buffer episodico che consente di riconoscere informazioni già note. Infine la formulazione implica l'elaborazione e produzione di informazioni visive e verbali, che necessita dell'armonica interazione tra tutte le parti coinvolte. Inoltre, un gran numero di fattori extralinguistici e particolari influenzano il processo e necessitano di essere gestiti, aggiungendo complessità ad un'attività che richiede già tutte le capacità mnesiche, organizzative e di concentrazione del cervello.

Come si è detto, la capacità della MdL è limitata nella quantità di informazioni che può gestire e nel tempo di ritenzione di tali informazioni. Queste limitate capacità influenzano tutte le operazioni mentali, come il processo traduttivo (Kosma, 2007); la facilità con la quale le informazioni vengono dimenticate comporterà che ogni volta che si recepiscano nuove informazioni, quelle precedenti vadano perdute, come quando iniziando a tradurre un nuovo segmento il traduttore non sarà in grado di tenere a mente

i segmenti già tradotti. Questo è fondamentalmente ciò che spinge i traduttori a tornare indietro nel testo per verificare le soluzioni già adottate, dato che le nuove informazioni fungono da interferenza e causano il decadimento delle informazioni precedentemente elaborate.

I limiti della MdL sono fondamentali anche per definire la quantità di informazioni testuali che la mente può elaborare insieme, riconducendoci con una prospettiva leggermente diversa alla questione discussa al paragrafo precedente (cf. 3.3) circa l'identificazione dell'unità di traduzione. Molti studiosi già negli anni '50 del secolo scorso hanno tentato di definire il numero massimo di elementi che la MdL riesce a trattenere insieme, identificando tale numero come variabile da 2 a 7 (Miller 1956, Bradbent 1975, Crowder 1976, cit. in Dragsted 2004: 44). Si è inoltre visto che quando tali elementi si richiamano tra di loro, la memorizzazione risulterà più facile ed essi formeranno un *chunk*, un blocco di informazioni interdipendenti che vengono memorizzate insieme. Tuttavia, vi è una differenza tra la semplice attività mnemonica e lo svolgimento di un'attività cognitiva complessa, e gli studiosi riconoscono che nel secondo caso i dati che la memoria riesce ad elaborare e a trattenere sono persino minori.

### **3.4.3. La memoria di lavoro a lungo termine**

Ericsson e Kintsch (1995) ipotizzano che, considerata la grande mole di informazioni che la MdL deve gestire nello svolgimento delle attività cognitive complesse e specializzate, debba esistere una sorta di magazzino intermedio tra la MdL e la MLT, nel quale sia possibile salvare tutti i dati utili in modo da averne una fruizione rapida durante il lavoro. La *memoria di lavoro a lungo termine* – così viene denominato tale nuovo spazio mentale – metterebbe così insieme la capienza della MLT e la capacità di gestione della MdL per rendere possibile lo svolgimento di quei compiti che per loro natura implicano l'elaborazione rapida e continua di grandi quantità di dati ed informazioni. La MLT, infatti, pur possedendo uno spazio illimitato, implica tuttavia dei tempi di evocazione troppo lunghi per questo genere di attività, quindi attraverso la memoria di lavoro a lungo termine (MdLLT) il cervello sfrutta solo la capacità spaziale della MLT, e la utilizza come estensione della MdL. Ericsson e Kintsch (1995) hanno dimostrato che nelle attività specializzate l'immagazzinamento delle informazioni avviene spesso in maniera inconscia, e che la mente di chi è solito svolgere tali attività è

in grado di riconoscere e selezionare in maniera automatica le informazioni che saranno rilevanti nel corso del lavoro, anticipando quindi eventuali difficoltà future premunendosi dei dati necessari ad appianarle.

Grazie al rapido accesso a grandi quantità di informazioni tramite la MdLLT, i traduttori professionisti che svolgono un'attività cognitiva molto complessa possono tradurre velocemente e gestire TU più lunghe e strutturate. Indubbiamente ciò sarà anche frutto della loro esperienza e del bagaglio di conoscenze e competenze acquisito; lo sviluppo della MdLLT può essere inteso come una capacità inconsapevole che consente di accelerare i ritmi di lavoro e gestire le informazioni in maniera funzionale accrescendo la grandezza delle TU, potendo usufruire di una "scorciatoia" per l'immagazzinamento e il recupero di dati dalla MLT.

Volendo riportare quanto sopra esposto al nostro principale argomento, ovvero l'adozione dei CAT Tools nel processo traduttivo, è interessante compiere un confronto tra alcune funzioni della memoria umana e le funzionalità delle memorie di traduzione, al fine di comprendere meglio, pur coscienti delle dovute differenze, come queste ultime siano state sviluppate a partire dalle prime. Innanzi tutto, la divisione tra diversi tipi di memoria si può facilmente applicare anche alle memorie digitali, specialmente attuando il paragone col più complesso sistema di gestione delle memorie adottato dalle agenzie di traduzione. Si può infatti associare la MLT ad una TM *server-based*, ovvero la memoria madre della quale l'agenzia invia copie ai suoi collaboratori, e che viene accresciuta con nuove informazioni solo dopo che esse siano state validate ripetutamente da un utilizzo condiviso dai vari traduttori e revisori, così come la MLT acquisisce nuove informazioni solo dopo che esse sono state reiterate e quindi confermate diverse volte. Ogni traduttore lavora con una TM derivante dalla memoria madre ma modificabile e contenente principalmente le informazioni utili allo svolgimento della traduzione assegnata. Questa TM più specifica può essere associata alla MdLLT, essendo di capacità comunque illimitata, di immediato utilizzo e contenente informazioni necessarie al compito da svolgere. Anche nel caso dei traduttori che lavorano autonomamente, una buona prassi spesso adottata è quella di utilizzare le TM ufficiali solo come consultazione, quindi senza modificarle, raccogliendo tutto il nuovo bagaglio di informazioni in nuove TM. Alla fine del lavoro, quando tutti i dati saranno confermati, essi verranno definitivamente acquisiti nella TM principale. Infine, la MdL, capace di gestire ed

elaborare in maniera consapevole i diversi input ricevuti al fine di produrre un risultato, può essere identificata col Workbench stesso, ovvero l'ambiente di lavoro nel quale confluiscono tutti i contributi dei vari strumenti utilizzati (TM, glossario, traduzione automatica, *Concordancer*, testo visualizzato nella schermata di lavoro ecc.), analizzati ed elaborati per produrre la traduzione. Tuttavia, mentre nella MdL le informazioni sono suscettibili di un rapido decadimento, l'ambiente di lavoro del CAT Tool consente la visualizzazione prolungata degli input. Anche l'evocazione delle informazioni è simile, essendo i dati della TM estratti in maniera automatica grazie alla somministrazione di uno stimolo, allo stesso modo della MLT che evoca più o meno volontariamente il ricordo sulla base di una sollecitazione. I tempi di recupero ovviamente sono diversi, quelli naturali sono più lunghi, ma la loro riduzione a vantaggio della produttività è proprio una delle finalità dell'adozione degli strumenti tecnologici.

#### **3.4.4. L'influenza dei CAT Tools**

Che tipo di influenza esercita sul processo mnemonico l'utilizzo dei CAT Tools? Purtroppo, come più volte puntualizzato, gli studi teorici ed empirici a riguardo non sono numerosi, soprattutto in quest'ambito così difficile da osservare. Tutto ciò che è possibile fare è formulare ipotesi sulla base delle teorie e degli studi disponibili. Prima di tutto, l'adozione di un sistema di gestione del lavoro quale è a tutti gli effetti una piattaforma CAT, richiede il ricorso ad un tipo specifico di memoria, ovvero quella procedurale, che implica il saper compiere un'azione in maniera automatica, senza doverci pensare, come ad esempio guidare l'auto o nuotare. È importante che l'uso del software non distolga la concentrazione dal compito principale, che è quello di tradurre, altrimenti il vantaggio sperato con la sua adozione si trasformerebbe in svantaggio.

Dagli studi già citati in questo capitolo (Dragsted 2004, Christensen 2011, O'Brien 2010 etc.), in particolare da quanto affermato nei questionari valutativi dai partecipanti agli esperimenti, è possibile dedurre che i rapidi suggerimenti delle TM e dei glossari, nonché la consapevolezza che ogni passaggio del lavoro venga registrato e recuperato con estrema facilità, inibiscono la messa in funzione sia delle capacità di evocazione delle informazioni possedute, sia di memorizzazione di quelle nuove (cosa che, tuttavia, può avvenire in maniera involontaria nel caso in cui vi sia un alto tasso di ripetitività). Diversi tra i partecipanti si dicono consapevoli di questa influenza e nella maggior parte dei casi la valutano in maniera negativa, come una mancanza di "controllo" sul processo. In realtà,

essendo l'aspetto mnesico legato principalmente alla disponibilità delle informazioni necessarie, in particolare della componente terminologica, il contributo del software si può valutare in termini positivi, poiché esso consisterà prevalentemente nell'accelerazione dei tempi di recupero di tali informazioni. La possibilità di accettare o respingere i suggerimenti ottenuti dipenderà comunque dalle scelte del traduttore, ed è già stata discussa nel capitolo precedente (cf. 3.3). Inoltre, il possibile impatto sulla minore memorizzazione della terminologia da parte del traduttore dovrebbe essere testata in maniera empirica per comprenderne l'entità. Ad ogni modo, anche nel caso in cui tale osservazione desse esito negativo, ovvero qualora i CAT Tools non incidessero sulla capacità di ricordare o meno i dati acquisiti, occorre tenere presente il naturale decadimento di tali informazioni sul lungo periodo, decadimento al quale le memorie di traduzione digitali non sono soggette.

Un'ultima considerazione interessante ci viene fornita da Pym (2011a), il quale afferma che la memoria è la dimensione del processo traduttivo che più viene intaccata dalla tecnologia, essendo esternalizzata ed estesa potenzialmente all'infinito. Osservando il processo di traduzione fondamentalmente come un'attività che implica la ricerca di soluzioni a problemi linguistici, più in particolare la ricerca di espressioni nella lingua di arrivo equivalenti a quelle della lingua di partenza, Pym vede nell'abbondanza di soluzioni proposte dalle risorse tecnologiche un elemento che può rendere più complessa l'identificazione della migliore soluzione da adottare. Egli puntualizza il rischio effettivo che la tecnologia non renda sempre il lavoro più semplice, ma che possa anche complicarlo e allungarne i tempi, per far fronte alla maggiore mole di informazioni da analizzare.

### *3.5 Pause come indicatori di attività cognitiva*

Concludiamo questa lunga discussione sugli effetti profondi dell'adozione dei CAT Tools guardando ad un aspetto che è stato considerato molto significativo dagli studiosi (Schilperoord 1996; Krings 2001; Dragsted 2004 e 2006; O'Brien 2006), ovvero la presenza delle pause e la loro importanza in quanto rilevatori di attività cognitive in corso. È opinione comune che, durante una pausa nell'utilizzo della tastiera per scrivere, o nella fissazione dello sguardo su un dato punto dello schermo, la mente del traduttore sia impegnata a svolgere attività cognitive che richiedono la sua concentrazione. Secondo

Schilperoord (1996: 9), “*pauses in text production are behavioral reflections of the cognitive processes involved in changing attentional state*”. Ciò significa che al sollevarsi di un problema o di un’incertezza, il livello di attenzione si innalza, e l’attività viene sospesa per analizzare la circostanza e cercarne la soluzione, o per evocare dalla memoria a lungo termine le informazioni utili al superamento della criticità.

### **3.5.1. Funzione delle pause nel processo traduttivo**

Le pause possono avere una duplice utilità, quella di programmare il lavoro successivo e quella di trovare soluzioni ai problemi che insorgono. Spesso le pause si posizionano in corrispondenza di categorie testuali ben definite (Séguinot cit. in O’Brien 2006), come all’inizio di una TU, e precedono la fase di produzione. È quindi possibile prevedere in che punti di un testo sia più probabile che avvenga una pausa “funzionale”. Infatti, è durante la pausa che la traduzione della TU inizia a prendere forma nella mente del traduttore, grazie all’elaborazione da parte della memoria di lavoro di tutte le informazioni tratte dal testo e recuperate dalla memoria a lungo termine. Possono però esserci pause finalizzate a rivedere quanto appena prodotto per valutarne la correttezza di forma e contenuto, o comunque è possibile e sensato sostenere che una stessa pausa, situata tra due TU, serva a svolgere entrambi i compiti simultaneamente o a rileggere l’ultima parte di testo tradotto, per poter proseguire in continuità con quanto già prodotto. Inoltre, sempre Schilperoord (1996: 89) afferma che vi sono livelli diversi di pause, più o meno lunghe in base alla loro localizzazione all’interno del testo. Le loro proporzioni rispecchiano la lunghezza del testo da rileggere o valutare, quindi più prolungate per paragrafi e frasi, più brevi all’interno delle frasi stesse, tra le parole che le compongono. Per l’autore, tuttavia, la durata delle pause non ne determina una gerarchia in termini di importanza, quanto piuttosto è indicatore di una maggiore difficoltà nel recupero delle informazioni necessarie ai diversi livelli testuali coinvolti, o può essere segno dello svolgimento di attività cognitive diverse – principalmente concettuali e semantiche quelle più lunghe, marcatamente linguistiche quelle più brevi. La necessità di pianificare il lavoro da svolgere su porzioni di testo più o meno grandi è un’altra possibile causa della differenza nella durata delle pause; la pianificazione per porzioni più ampie richiederà nella maggior parte dei casi tempi più lunghi che la pianificazione a livello lessicale. In generale, sostiene Dragsted (2004: 74), le pause associate ai livelli testuali inferiori sono più significative dal punto di vista della traduzione, trattandosi in molti casi di pause tra

TU. Esse saranno funzionali sia alla formulazione di scelte lessicali e sintattiche, sia all'organizzazione del lavoro sebbene a livello di frase, cosa che rende la traduzione diversa dalle altre attività di produzione testuale.

La lunghezza delle pause varia anche in base alle difficoltà incontrate nel produrre l'unità di testo successiva. In generale, dover affrontare una problematicità richiede una pausa più lunga di quella richiesta per la semplice pianificazione del lavoro, poiché le informazioni da recuperare dalla memoria a lungo termine sono nel primo caso più difficilmente accessibili. La difficoltà è un fenomeno cognitivo, e affrontarla impegna considerevolmente le capacità della memoria di lavoro, molto più che nell'elaborazione di porzioni di testo non problematiche, per questo la mente ha bisogno di fare una pausa e concentrare tutte le sue capacità sulla problematicità incontrata. Secondo Krings (2001), le pause che durano più di tre secondi indicano la presenza di un problema, e testi che presentano diverse problematicità implicano tempi di produzione più lenti e prolungati.

### **3.5.2. L'impatto dei CAT Tools**

Negli studi sui CAT Tools, si è riscontrato che l'adozione di una TM e la suddivisione in segmenti corrispondenti più o meno al livello delle frasi, alteri l'abitudine dei traduttori relativamente alle pause nel corso del lavoro. Prima di tutto, occorre dire che dai diversi studi si è evidenziato che ogni traduttore ha un proprio ritmo di lavoro, posiziona le proprie pause in momenti e a livelli confacenti al proprio metodo di lavoro, e può produrre pause di durate diverse, senza che ciò alteri significativamente la qualità del lavoro finale (Dragsted 2004: 167; O'Brien 2006: 5). Quindi, non esiste uno schema comportamentale condiviso dei tempi e delle pause nel lavoro, e tale variabile è esclusivamente dipendente dall'individuo e dalla propria prassi lavorativa. Inoltre, con l'aiuto dei già citati software di registrazione del comportamento al computer (utilizzo di tastiera e mouse, e in alcuni casi anche il rilevamento del movimento oculare), gli studi hanno messo in luce la tendenza di molti traduttori, specialmente quelli professionisti, a non fermarsi prima di nuove frasi o paragrafi, o comunque a non fare delle pause più lunghe che in altre parti del testo. L'utilizzo di un CAT Tool modifica questa tendenza, data la necessità di effettuare delle azioni pratiche di gestione del software come immettere i comandi che servono a passare da un segmento all'altro o confermare il segmento precedentemente tradotto. Tali azioni impongono al traduttore di interrompere il processo di traduzione e si presume che implicano anche un momento di rapido

recupero delle informazioni appena prodotte e di quelle da elaborare, per ravvivare la memoria a breve termine dopo l'interferenza rappresentata dalle operazioni richieste dal programma.

Si è già detto che i traduttori professionisti sembrano in generale non avere di bisogno di pause prima di iniziare un nuovo segmento. Quando ciò avviene, la pausa è significativa perché implica un'attività di *problem-solving* in corso, come la necessità di modificare la struttura sintattica della frase di partenza, ampliando o restringendo il segmento di testo con operazioni che di nuovo richiedono l'esecuzione di alcuni comandi del programma, ma che principalmente implicano un momento di valutazione e decisione che necessita di una pausa apprezzabile. Secondo i dati raccolti da Dragsted (2004), l'utilizzo di un CAT Tool introdurrebbe per i traduttori la necessità di fare una pausa all'inizio di ogni TU, con un incremento di tale abitudine di circa il 20% per i professionisti, e di circa il 4% per gli studenti, che diversamente dai primi dimostrano un comportamento prediligente le pause anche in assenza di CAT Tools. Ciò sembra rispecchiare anche quanto osservato relativamente alla differenza dei tempi di produzione e revisione tra studenti e professionisti. Questi ultimi vanno più spediti nella fase di produzione, per poi soffermarsi maggiormente durante la revisione, mentre gli studenti hanno un comportamento esattamente inverso, dedicando maggior tempo alla prima fase di produzione e cercando di risolvere a monte le difficoltà che incontrano, destinando poi alla revisione un tempo molto più breve. Ciò si tradurrà per i professionisti in poche e brevi pause nella fase di produzione ma molto più numerose e prolungate in fase di revisione, ed esattamente il contrario per gli studenti.

Tornando ai CAT Tools, abbiamo detto che essi accrescono apprezzabilmente il numero delle pause ad inizio segmento per i traduttori professionisti, e meno significativamente per gli studenti. Ciò può essere anche sintomo di un diverso rapporto con lo strumento digitale, in quanto i professionisti ritengono necessario accertarsi che i suggerimenti delle TM siano corretti prima di adottarli, riponendo nel software una limitata fiducia che necessita sempre del vaglio e possibilmente dell'intervento da parte del traduttore. Come già puntualizzato, gli studenti tendono invece ad adottare pressoché indiscriminatamente le soluzioni derivate dalla TM. Dalla verbalizzazione o dai questionari successivi al processo, emerge tra l'altro una diversa percezione delle pause. Gli studenti ne sono meno consapevoli, proprio perché inclini in ogni caso a compierle

anche nei processi traduttivi classici e per l'utilizzo spesso poco critico del CAT Tool. I traduttori professionisti invece lamentano l'accresciuta discontinuità causata dal software, la frammentazione del lavoro e l'allungamento dei tempi di traduzione, riconoscendo che il nuovo comportamento indotto dall'adozione del software con l'aumento del numero di pause in fase di produzione implichi lo spostamento di alcune operazioni cognitive solitamente attivate in fase di revisione da quest'ultima alla fase di produzione, con una conseguente diminuzione dei tempi di revisione apprezzata durante gli esperimenti di raccolta dei dati. Tutto questo implica come ipotizzato una significativa modifica del processo traduttivo e delle abitudini operative e produttive.

### *3.6 Conclusioni*

Prima di procedere oltre con la trattazione, è utile sintetizzare brevemente le principali conclusioni alle quali si è addivenuti. Pur essendo queste spesso mancanti di solide basi di osservazione empirica, o mero frutto di ragionamenti astratti e per ipotesi, gli studi ai quali si è fatto riferimento, e l'esperienza maturata personalmente lasciano credere che esse non si discostino molto dal reale.

1. I CAT Tools sono ormai divenuti strumenti indispensabili per il lavoro di traduzione tecnica, specialmente per quei linguaggi specialistici che presentano un elevato tasso di ripetitività e necessitano di un perfetto controllo della terminologia utilizzata;

2. L'utilizzo dei CAT Tools risulta più congeniale a chi li adotta come strumenti di lavoro fin dall'inizio della propria carriera, mentre si nota una certa riluttanza nei traduttori che per anni non ne hanno fatto uso, a causa della necessità di investire tempo nell'apprendimento di una nuova modalità operativa che può in un primo momento causare una diminuzione dei ritmi di lavoro e quindi della produttività;

3. I CAT Tools modificano alcuni aspetti del processo traduttivo, imponendo di lavorare secondo dei principi o delle prassi che spesso non corrispondono a quelle naturali dei traduttori. Le modifiche, prima di essere pratiche, sono di natura cognitiva, riguardanti il modo in cui il traduttore percepisce il proprio lavoro e la propria figura professionale, ma anche il testo da tradurre e il prodotto finale. In particolare:

- a. Cambia l'ambiente di lavoro, con la possibilità di avere tutti gli strumenti e le risorse di riferimento visualizzabili nella stessa schermata in cui si traduce, cosa

che ne favorisce la fruizione ma che richiede anche capacità di concentrazione e di mantenimento di una certa “autonomia operativa” dai suggerimenti ricevuti da TM e glossari;

- b. Si può notare che i suggerimenti della TM esercitano un’influenza più forte negli studenti piuttosto che nei traduttori professionisti, essendo spesso adottati incondizionatamente dai primi, e valutati criticamente dai secondi. Tuttavia, anche nei traduttori professionisti e specialmente in condizioni di lavoro particolari la tendenza a “dare per buoni” gli exact match prodotti dalle memorie digitali fa sì che parte del lavoro venga delegato acriticamente alla macchina. Ciò causa un certo fastidio ed un senso di frustrazione, specialmente quando la richiesta di attenersi alle versioni in memoria non è una scelta del traduttore ma un’imposizione da parte dell’agenzia o del cliente finale;
- c. Viene modificato l’approccio personale al testo, con la nuova prassi di lavoro che prevede la segmentazione in frasi. Tale divisione non corrisponde in molti casi alla determinazione naturale delle TU cognitive, spesso più piccole in lunghezza. Ciò comporta anche la perdita di una visione d’insieme del testo e il rischio di elaborazione disarmonica delle sue parti, data la segmentazione imposta e spesso la divisione dello stesso testo tra più traduttori. La sensazione che di questo aspetto hanno i traduttori è negativa relativamente al rischio di perdere importanti informazioni contestuali, e alla necessità di adattare la propria procedura operativa a quella più rigida e per certi versi innaturale imposta dal CAT Tool. Tuttavia, si riconosce anche che la presentazione del testo frase per frase favorisce la concentrazione e riduce la tendenza a “vagare” lungo l’interno documento.
- d. Fare affidamento sulle memorie digitali implica utilizzare meno le capacità mnesiche proprie del traduttore. Ciò può causare un senso di frustrazione nel traduttore, ma al contempo gli consente di adoperare le proprie capacità mentali per aspetti diversi dal recupero mentale della terminologia nota. In ogni caso, le TM garantiscono l’assenza di decadimento sul lungo periodo e consentono un recupero più rapido delle informazioni rispetto alle capacità della memoria a lungo termine. Si è fatta propria l’ipotesi della creazione di una memoria di lavoro a lungo termine, capace di gestire velocemente grandi quantità di informazioni senza dover contare sui lunghi tempi di recupero della MLT.

- e. Cambia anche il comportamento relativo alle pause, indicatrici di attività cognitive in corso, attività che possono essere di problem-solving o di pianificazione del lavoro. I traduttori professionisti che non sono abituati a soffermarsi tra una frase e un'altra ma che al contrario preferiscono procedere velocemente durante la fase di produzione della traduzione si sentono invece costretti a fermarsi a causa della divisione in segmenti e dei comandi del software che occorre attivare per passare da un segmento all'altro. Ciò comporta per loro un allungamento dei tempi di produzione e lo spostamento di alcune attività cognitive dalla fase di revisione a quella di produzione. Viene inoltre avvertita negativamente la discontinuità nell'elaborazione testuale e la necessità di interrompere un'attività complessa come la traduzione con attività funzionali e meccaniche come i comandi del CAT Tool. Pur divenendo essi col tempo automatici, impongono un ritmo diverso da quello normalmente sviluppato dai traduttori professionisti.

## Capitolo 4: CAT Tools strumento e oggetto della didattica

### 4.1 *Apprendere una competenza professionalizzante*

#### 4.1.1. **La dimensione didattica**

Chiudiamo questa Prima parte dedicando un capitolo alla didattica della traduzione in ambito universitario, e a come i CAT Tools vi si inseriscano come strumenti per una formazione professionale. Da quando il loro utilizzo è divenuto un requisito fondamentale per proporsi sul mercato della traduzione, gli atenei hanno cominciato a inserire, nel piano formativo dei curricula in traduzione, uno o più corsi o laboratori per l'apprendimento dei CAT Tools e delle altre tecnologie di sussidio alla traduzione. Come puntualizzano Olvera *et al.* (2005: 132), data la natura multidisciplinare della traduzione e la diversità di competenze che essa richiede, i programmi formativi devono offrire non solo una preparazione linguistica, ma conoscenze di pari importanza relative alle diverse discipline coinvolte nel processo traduttivo. Calzada Perez (2005) sottolinea che proprio l'integrazione tra teoria e pratica all'interno della didattica della traduzione è stata la chiave per la grande crescita (quantitativa e qualitativa) che i corsi accademici in traduzione hanno conosciuto nell'ultimo ventennio. Un corso di traduzione completo dovrebbe fornire conoscenze di teoria e pratica della traduzione, padronanza dei linguaggi specialistici e competenze pratiche come le tecniche di archiviazione e organizzazione della terminologia e delle risorse di riferimento. Essendo la competenza tecnologica e nello specifico la padronanza dei CAT Tools un requisito essenziale per il mercato, questo dovrà trovare il giusto spazio all'interno della didattica, per consentire agli studenti di ottenere una buona conoscenza di tutti gli strumenti digitali applicati all'ambito della traduzione (Olvera *et al.* 2005: 136).

Con la progressiva affermazione dei corsi pratici di sussidi per la traduzione nei curricula accademici, si è andata anche sviluppando una nuova teoria della didattica, che ha tenuto conto dell'apporto di tali strumenti, delle nuove esigenze poste e delle diverse prospettive

aperte rispetto alla didattica classica della traduzione. Infatti, già il dover riconoscere la necessità di fornire agli studenti una conoscenza pratica di queste risorse equivale a riconoscere un cambiamento nel panorama della traduzione come attività e oggetto di studio: l'introduzione di un approccio fondamentalmente funzionale che tenga in considerazione una serie di aspetti determinati da motivazioni e fenomeni non strettamente linguistici, come il mercato e le ragioni economiche che lo determinano. Oggi è universalmente riconosciuta una diversità di metodo e stile tra la traduzione letteraria e quella tecnico-scientifica, con corsi specifici dedicati a questi due ambiti nella maggior parte degli atenei, al fine di fornire agli studenti i fondamenti necessari tanto per l'uno quanto per l'altro ambito in termini di conoscenza delle tipologie testuali coinvolte e delle più adeguate strategie di traduzione applicabili<sup>21</sup>. Per i corsi di traduzione letteraria, ad esempio, si dà maggiore peso ai laboratori di scrittura creativa, allo studio dei diversi stili letterari e allo sviluppo delle capacità di riproduzione di questi stili in traduzione. I corsi dedicati al ramo tecnico-scientifico, invece, dedicano maggiore attenzione allo studio dello *specialized discourse*, delle caratteristiche dei maggiori linguaggi specialistici e allo sviluppo di una capacità di scrittura che sappia riprodurre tali caratteristiche e applicare le tecniche e le strategie più consone. In alcuni casi, tuttavia, ciò incontra il dissenso dei teorici della traduzione che fanno propria una prospettiva traduttologica (Ladmiral, 2004), marcatamente teorica, che osserva la disciplina come un unico fenomeno governato da leggi e dinamiche uniche malgrado le sue numerose sfaccettature e articolazioni. Essi temono che l'applicazione selettiva dei CAT Tools a determinate tipologie testuali possa lasciare intendere che esistano differenze di sostanza all'interno del processo traduttivo per testi letterari e per testi tecnici, e nella lingua utilizzata per l'uno o per l'altro scopo, come se si venisse a creare una scissione all'interno della disciplina, o una netta separazione all'interno della lingua stessa. Occorre puntualizzare che il riconoscimento di queste diversità tra i due ambiti non presuppone una divisione sostanziale all'interno degli studi sulla traduzione, la cui natura epistemologica rimane unica, come uniche sono le teorie che ad essa fanno riferimento. In realtà, riconoscendo i CAT Tools come strumenti atti ad agevolare alcuni processi all'interno del complesso processo traduttivo, il fatto che essi si applichino con

---

<sup>21</sup> Si vedano le offerte formative delle principali università italiane che offrono corsi in traduzione, come l'Alma Mater di Bologna (<http://www.scuolalingue.unibo.it/it>) o la Scuola superiore per interpreti e traduttori di Trieste (<http://iuslit.units.it/>)

maggior utilità ad alcune tipologie testuali e ambiti linguistici piuttosto che ad altri non sta a significare altro che una maggior presenza nei primi di quelle caratteristiche formali, stilistiche e sintattiche che ben si coniugano con le capacità proprie di questi strumenti, e che si collegano alle finalità perseguite da tali testi.

#### **4.1.2. L'approccio costruttivista e la centralità del laboratorio**

L'adozione dei CAT Tools, così come di ogni altro strumento di uso pratico nella didattica, ben si sposa con la teoria costruttivista (Danks, 1997: 154) che afferma che la conoscenza umana è frutto dell'interazione tra le esperienze e le idee che tali esperienze suscitano. La riflessione sull'esperienza vissuta produce uno schema comportamentale che con la pratica e la riflessione continuata diventa conoscenza. L'approccio costruttivista incentiva l'esperienza didattica pratica, il così detto *learning by doing*, l'adozione di materiale di uso reale, la collaborazione e la sperimentazione, seguendo un percorso di apprendimento deduttivo invece che induttivo. Il docente guida il percorso formativo, lasciando però ampio spazio all'azione individuale degli studenti. I CAT Tools presuppongono un utilizzo personale ma al contempo uno scambio collaborativo (come nella condivisione delle TM e dei glossari, o nella divisione del lavoro tra i membri di un team); il loro utilizzo è essenzialmente pratico e non può essere appreso se non tramite esperienza diretta. L'ambiente più adatto alla didattica dei CAT Tools è il laboratorio (Kenny e Way, 2001: 14; Doherty e Morkens, 2013: 122; Li, 2002: 514) in cui generalmente i corsi vengono organizzati secondo una logica graduale, partendo dalle funzioni basilari per giungere infine alla padronanza dello strumento nelle sue funzioni più avanzate. La gradualità riguarda anche la tipologia di testi che vengono utilizzati, più semplici e con grado elevato di ripetitività agli inizi per acquisire praticità con il programma, più complessi e con un maggior focus sulla terminologia in seguito, quando l'utilizzo del software sarà già stato acquisito e si passerà a studiarne l'applicazione in ambito terminologico e linguistico. All'interno del laboratorio, in cui i computer sono spesso messi in rete, è possibile sperimentare la dimensione collaborativa, simulando una vera situazione di lavoro in cui un testo viene suddiviso fra più traduttori che utilizzano le stesse risorse di riferimento. In alcuni casi, nelle esercitazioni è persino possibile suddividere la classe in traduttori e revisori, per consentire agli studenti di correggersi reciprocamente e apprendere dagli errori (così come dalle frasi correttamente tradotte) e dalle revisioni dei colleghi, facendo allo stesso tempo esperienza di come si svolge

normalmente il lavoro del traduttore o del revisore. Una prospettiva didattica lungimirante dovrebbe infatti formare gli studenti al fine di far loro comprendere le principali mansioni delle diverse figure professionali all'interno delle agenzie di traduzione, dal traduttore al project manager, al curatore della terminologia fino al revisore, avendo chiari i compiti demandati ad esse, e il modo in cui esse collaborano e interagiscono.

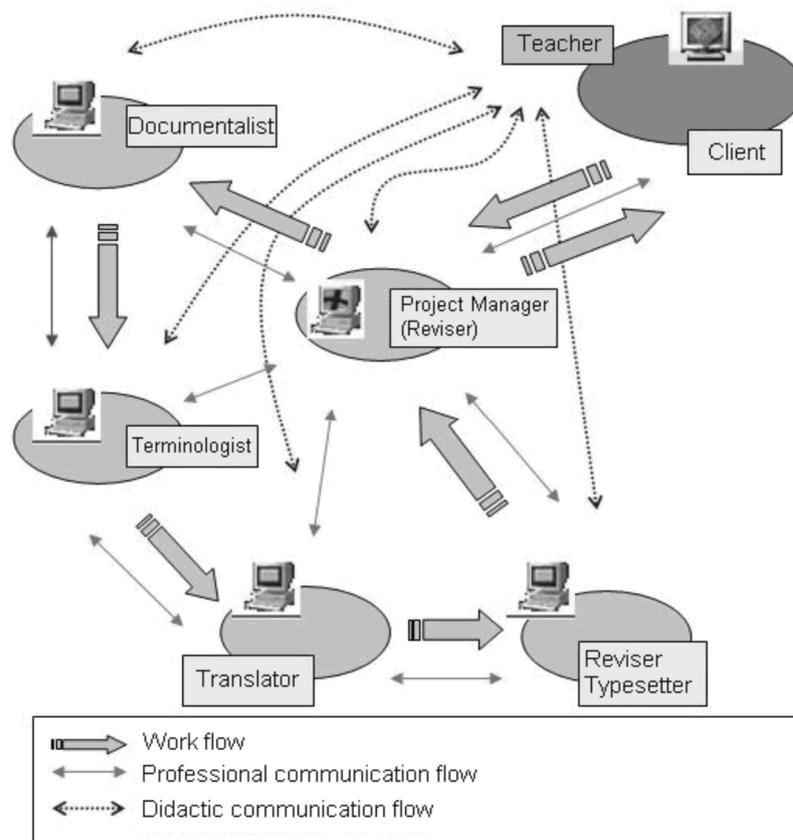


Fig. 26: Le interazioni tra i diversi attori nel processo traduttivo (Olvera et al. 2005: 138)

Il progetto didattico “Aula.int”<sup>22</sup> sviluppato dalla Facoltà di Traduzione ed Interpretariato dell’Università di Granada (Olvera *et al.* 2005: 138), prima sperimentazione dell’approccio professionale alla didattica della traduzione, è un interessante esempio di riproduzione in una classe virtuale delle dinamiche lavorative tipiche delle agenzie di traduzione. Il suo scopo è quello di rendere gli studenti consapevoli delle reali condizioni di lavoro a livello di organizzazione, divisione dei ruoli e delle responsabilità, procedure

<sup>22</sup> <http://aulaint.ugr.es>

e relazioni. L'immagine sopra riportata rappresenta i principali attori partecipanti al mercato della traduzione, uniti da diverse frecce. Come riportato nella leggenda in basso, le frecce più grandi mostrano i passaggi della traduzione dal cliente passando per le varie figure professionali coinvolte, per tornare infine al cliente; le frecce sottili continue rappresentano la rete di comunicazioni tra i diversi professionisti, infine le frecce tratteggiate indicano le comunicazioni didattiche, ovvero quelle tra il docente – che nella simulazione ricopre il ruolo del cliente – e gli allievi che ricoprono i diversi ruoli come all'interno di un'agenzia. La metodologia di simulazione dell'ambiente professionale reale si è dimostrata efficiente nell'incoraggiare gli studenti alla collaborazione, alla partecipazione attiva e ad una riflessione critica e cosciente sulle varie fasi del processo traduttivo che essi stessi hanno potuto sperimentare, nonché nello sviluppo di un approccio eclettico e flessibile che tenesse conto dell'articolata realtà del lavoro nel mercato della traduzione (*ibid.*).

I corsi di apprendimento dei CAT Tools non sono esclusivamente indirizzati a studenti. Esistono corsi, dentro e fuori i contesti universitari, destinati a professionisti che intendano investire sulla propria formazione permanente. A tal proposito, considerando le differenze nell'approccio di studenti alla prima esperienza e professionisti già inseriti nel settore, è possibile prendere in prestito dalla teoria costruttivista due termini che identifichino i diversi processi di apprendimento messi in atto. Gli studenti alla prima esperienza sia con i CAT Tools sia con la traduzione hanno uno schema mentale in formazione; la nuova esperienza si inserirà all'interno di tale schema contribuendo alla sua creazione con le rappresentazioni delle nuove conoscenze acquisite. Tale processo viene detto “assimilazione”, mentre nel secondo caso, quello di professionisti che intendono utilizzare un CAT Tool dopo anni di lavoro svolto senza, si parla di “conciliazione<sup>23</sup>”, perché avendo l'individuo uno schema mentale già formato ed escludente il ricorso ai CAT Tools, questi rappresenteranno un nuovo elemento che apporterà un cambiamento allo schema mentale preesistente, implicherà una negoziazione tra vecchie e nuove conoscenze e il ripristino di un nuovo equilibrio tra di esse.

---

<sup>23</sup> I termini originari in inglese sono *assimilation* e *accomodation* (Danks, 1997: 156)

#### 4.1.3. L'apporto dei CAT Tools nella didattica

La discussione sull'importanza dell'insegnamento dei CAT tools nei corsi universitari di traduzione ha già alle spalle più di un decennio. Studiosi come Scherf (1992), Haynes (1998) e Kenny (1999) hanno più volte sottolineato il contributo di tali sussidi alla formazione del futuro traduttore. I CAT Tools inseriti all'interno di un corso di traduzione, se correttamente valorizzati, possono assumere una importante applicazione didattica, che come si vedrà più avanti, può apportare benefici sotto diversi aspetti. La formazione dovrà mostrare agli studenti sia gli aspetti positivi che quelli negativi di tali strumenti, educarli ad un uso critico di tali risorse e alla migliore applicazione delle funzioni e delle potenzialità che esse offrono. Se correttamente inquadrati, i CAT Tools potranno consentire agli studenti di comprendere più profondamente il processo traduttivo nella sua complessità; conoscere e padroneggiare il loro uso potrà diventare per i futuri traduttori uno strumento di professionalità e organizzazione del lavoro, di analisi testuale e studio della lingua, di ricerca terminologica e formazione settoriale, di creazione di un patrimonio linguistico che va ben al di là del testo tradotto.

Le possibilità offerte dall'introduzione delle tecnologie CAT nella didattica della traduzione sono così descritte da Kenny (1999: 78):

*By integrating CAT tools into our teaching environment, we are not merely imparting the kind of practical skill that will get graduates jobs. We are also creating an environment in which basic and applied research can be carried out in a number of areas, including translation pedagogy, terminography, CAT Tools evaluation, human-machine interaction, and text analysis and composition.*

L'apporto formativo è quindi ampio e multiforme. Bowker (2002: 15), riconoscendo inoltre che i CAT Tools apportano in ultima istanza delle modifiche alla natura e ai processi del lavoro di traduzione (cf. Capitolo 3), invita i formatori a tenere in considerazione la loro influenza e a renderne consapevoli gli studenti. Bowker sostiene che l'integrazione dei mezzi tecnologici nel curriculum accademico in traduzione possa produrre un impatto sul modo stesso in cui la traduzione viene insegnata, ad esempio con la possibilità di condividere risorse in maniera digitale, o di costruire poco per volta

modelli testuali e linguistici prodotti dagli studenti stessi e salvati nelle loro TM. Scherf (1992: 157) afferma che l'uso delle tecnologie consenta di sviluppare approcci individuali e personalizzati al processo didattico, grazie ai quali gli studenti possano sperimentare in prima persona e in maniera concreta l'ambiente e il processo traduttivo. I docenti possono a loro volta osservare in tempo reale i progressi degli studenti a confronto con il processo traduttivo, potendo dare a ciascuno indicazioni e consigli personalizzati, oltre alle spiegazioni di natura più generica destinate alla classe. Secondo Kenny (1999: 77) l'utilizzo dei CAT Tools induce gli studenti a considerare problematiche come la diversità delle tipologie testuali e le loro caratteristiche intra ed intertestuali che determinano se e in che misura tali tipologie possano beneficiare dell'applicazione delle memorie di traduzione. Inoltre, riflettendo sulla natura della macchina e sul modo in cui essa opera, gli studenti imparano a riconoscere il tipo di contributo che essa può fornire, e ad adattare le loro prassi operative al fine di massimizzare i benefici che possono trarre dall'utilizzo dei CAT Tools. Fondamentalmente, gli studenti divengono consapevoli del fatto che i computer non sono capaci di applicare ragionamenti intelligenti o conoscenze comuni allo stesso modo del cervello umano. Tale consapevolezza rende loro possibile gestire meglio l'utilizzo sia in vista dei risultati concretamente ottenibili, sia applicando le giuste tecniche per semplificare il lavoro al computer per ottimizzarne la performance. Comprendendo appieno le dinamiche di interazione uomo-macchina, gli studenti saranno inoltre in grado di valutare gli strumenti di cui si avvalgono, e scegliere anche, in base alle necessità, la risorsa più utile e più funzionale (Sager, 1994: 194). Gaspari (2001), riconoscendo la crescente importanza dell'integrazione della tecnologia nella didattica della traduzione, riflette sulla necessità di fornirne anche una conoscenza teorica, con nozioni di linguistica computazionale e applicata, *natural language processing*, intelligenze artificiali, etc. In tal modo gli studenti potranno più chiaramente comprendere le idee, le finalità, gli studi e le metodologie a partire dalle quali i principali strumenti digitali sono stati sviluppati, ed essere realmente consapevoli della natura di tali risorse, che influenzano in maniera diretta il lavoro del traduttore. Questa maggiore consapevolezza rende gli studenti capaci di saper valutare meglio gli aspetti positivi e le eventuali criticità presentate da ogni *tool*, non cadendo negli eccessi di fiducia o al contrario nella più negativa rinuncia di ogni ricorso ai sussidi tecnologici. Le risposte al questionario sviluppato da Gaspari (2001: 40-44) per captare le impressioni degli studenti

prima e dopo un corso incentrato sulla traduzione automatica e i CAT Tools hanno dimostrato come gli studenti manifestassero un iniziale scetticismo, dovuto alla mancanza di una reale conoscenza di tali strumenti. Alla fine del corso, l'atteggiamento preponderante era invece decisamente più bilanciato, con aperture positive e ottimistiche, scaturite dall'esperienza diretta all'interno dei laboratori pratici, dove gli studenti hanno potuto constatare con mano quanto appreso durante la parte teorica del corso.

Un utilizzo molto interessante che può essere fatto dei CAT Tools riguarda la loro capacità di raccogliere dati che possono essere utilizzati come risorsa nello studio della lingua nonché delle strategie traduttive. Tale aspetto, che verrà analizzato in dettaglio nel paragrafo seguente, può rivestire gli strumenti CAT di un potenziale di utilizzo decisamente superiore rispetto alla loro semplice applicazione come sussidi per la traduzione. Essi potranno essere utilizzati al pari di corpora elettronici bilingui e paralleli, all'interno dei quali osservare un utilizzo reale della lingua e in particolare dei linguaggi settoriali, traendone informazioni sulla terminologia, la sintassi, i costrutti più frequenti e lo stile.

## 4.2 *A sostegno dell'acquisizione di una competenza traduttiva*<sup>24</sup>

### 4.2.1. **In cerca di una definizione**

Lo scopo ultimo di qualsiasi corso in traduzione dovrebbe essere quello di fornire allo studente conoscenze ed esperienze che possano consentire di sviluppare una reale competenza traduttiva. Tutti gli insegnamenti, quelli teorici alla pari di quelli pratici, convergono in ultima istanza a formare un bagaglio di "saperi" che sono la base su cui intraprendere una carriera in ambito linguistico. Che cosa significa esattamente possedere una competenza nel campo della traduzione? Sebbene questo concetto possa sembrare abbastanza semplice – competenza in traduzione significa essenzialmente saper tradurre forma e contenuto di un testo da una lingua a un'altra – molti studiosi hanno cercato di darne una definizione, per comprenderne la natura. All'inizio la tendenza era quella di considerare la traduzione come un'attività che poteva essere svolta da chiunque possedesse una buona conoscenza di due lingue (Koller, 1979; Ballard 1984, cit. in Pym

---

<sup>24</sup> Sull'acquisizione della competenza traduttiva, è stato illuminante il saggio di Gregory Shreve (2006), a cui si rimanda per il confronto tra studenti ed esperti e sullo sviluppo delle strategie traduttive

2003: 491), quasi una capacità acquisita di pari passo con lo studio delle lingue, o posseduta in maniera innata nel caso di individui bilingui. Con l'avanzamento degli studi si affermò invece la natura peculiare di tale attività (PACTE, 2003: 45), richiedente un set di abilità non esclusivamente riconducibili alla padronanza di una o più lingue (Arnold, 2000: 1). È una competenza specialistica che mette insieme conoscenze dichiarative e procedurali (PACTE 2008: 106), e che si acquisisce con strategie e tecniche apposite. Wills (1976, cit. in Pym 2003) giunge persino ad affermare che la competenza traduttiva include un set di conoscenze talmente ampio da non poter essere chiaramente definibile. Ci sono studiosi, come Campbell (1991) che hanno cercato di determinare la competenza in traduzione in base alla valutazione di parametri diversi come la correttezza grammaticale, le scelte lessicali, l'adeguatezza della sintassi ecc., con complicatissime e in parte discutibilissime considerazioni sull'adeguatezza delle scelte compiute. Altri studiosi invece si sono cimentati in tentativi di definizione ricorrenti ora ad una logica sommativa di numerose sotto-competenze, ora ad una più sintetica descrizione per mettere in luce nel primo caso la natura multidisciplinare e variegata del processo traduttivo, o la caratteristica considerata peculiare nel secondo (Orozco e Hurtado Albir, 2002: 375-376). In generale, tuttavia, il criterio multi-componenziale è quello più diffuso. Uno schema tra i più noti per la rappresentazione della competenza in traduzione è quello sviluppato dal PACTE (2003), gruppo spagnolo di studio e ricerca sull'acquisizione della competenza in traduzione.

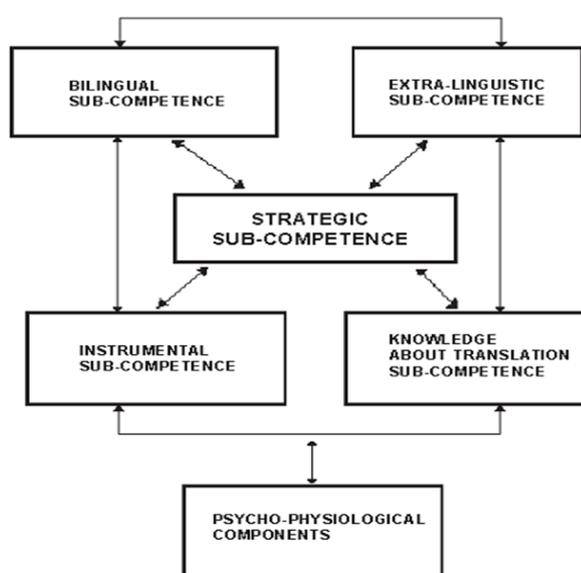


Fig. 27: Il modello di competenza traduttiva del PACTE (2003)

Questo schema è interessante non solo per il fatto che include già le competenze strumentali all'interno delle quali è possibile includere i CAT Tools e le altre risorse digitali per la traduzione. È significativo il fatto che le competenze strumentali vengano ricondotte alle strategie traduttive (riquadro centrale), sottintendendo che l'utilizzo di tali strumenti debba divenire strategico per il miglior svolgimento della traduzione. Ciò significa saper distinguere non solo le tipologie di strumenti più utili nelle diverse circostanze, ma anche quei casi in cui la loro adozione è in effetti la strategia più corretta da applicare. È interessante anche l'interrelazione resa evidente dalle doppie frecce tra le diverse abilità che formano la competenza in traduzione, perché in effetti ognuna è strettamente legata alle altre, pur trattandosi di competenze sia dichiarative (competenze extra-linguistiche e di conoscenza della traduzione) che procedurali (competenze linguistiche, strumentali e strategiche). Un ulteriore elemento di interesse sono le componenti psico-fisiologiche, ovvero quel set di attitudini innate come l'attenzione, la creatività, la memoria, l'approccio emotivo e le dinamiche cognitive soggiacenti all'intero processo, come suggerisce la posizione occupata dal riquadro corrispondente all'interno dello schema (PACTE, 2007: 97). La formazione in tutte le diverse aree di competenza identificate viene auspicata da Tan (2008) che riconosce la necessità di un approccio didattico integrato (*whole-person translation education*) per una reale competenza traduttiva.

Un altro schema che può rivelarsi interessante è quello pubblicato dall'EMT (European Masters in Translation, 2009, riportato in Pym 2013: 488):

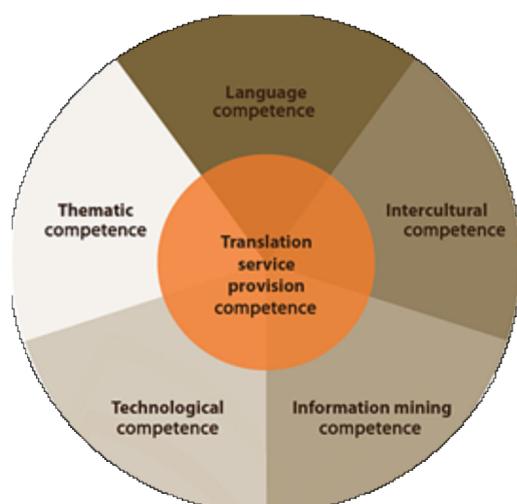


Fig. 28: Lo schema multi-componente della competenza traduttiva (EMT 2009)

Allo stesso modo dello schema formulato dal PACTE, anche questo include già le competenze tecnologiche, a controprova del fatto che esse sono ormai a tutti gli effetti considerate parte immancabile del bagaglio conoscitivo ed esperienziale di un traduttore. Prendendo come punto di partenza quest'ultimo schema, possiamo compiere una riflessione sull'impatto delle tecnologie per la traduzione in relazione alle altre competenze in esso riportate. Ad esempio, la capacità di estrarre informazioni (*information mining*) non è più una competenza a sé stante, poiché buona parte dell'estrazione avviene automaticamente dalle memorie, dai glossari o dalla traduzione automatica (TA). La capacità da sviluppare sarà piuttosto quella di capire quando e fino a che punto fidarsi dei risultati dell'estrazione automatica, e la capacità di selezionare tra le tante l'informazione più pertinente. Anche l'ambito delle competenze linguistiche (*language competence*) viene modificato in quanto molta terminologia e molte espressioni sono già presenti nelle memorie o suggerite (spesso anche in maniera erronea) dalla TA. Si potrebbe pensare che questo implichi un minore bisogno di formazione linguistica, ma al contrario essa continua ad essere indispensabile al fine di non cadere in errori dovuti ad eccessiva fiducia sull'affidabilità della macchina. Si potrebbe inoltre ipotizzare (Pym, 2013: 492) che l'introduzione delle tecnologie cambi il tipo di competenza linguistica necessaria ad eseguire una traduzione, poiché parte di essa può essere demandata ai software appositi. Si potrebbe giungere ad affermare che il traduttore non abbia bisogno di possedere una solida competenza della lingua di origine, essendo la traduzione vera e propria in buona parte prodotta da suggerimenti della TM o elaborazioni della TA. Basterebbe un esperto esclusivamente della lingua di arrivo e dell'area tematica del caso per svolgere un post-editing della traduzione prodotta dai sussidi digitali. La stessa logica va ad intaccare l'ambito delle competenze tematiche (*thematic competence*) potendo affermare che, poiché le TM/TA riducono la necessità di abilità linguistiche, al traduttore non è richiesto conoscere i linguaggi specialistici e le peculiarità sintattiche e terminologiche di un dato ambito settoriale, dato che tutte queste caratteristiche sono già gestite dai supporti digitali per la traduzione. Ovviamente si tratta di considerazioni per assurdo e portate all'estremo dall'autore, ma che dicono qualcosa di utile sulla necessità di inquadrare nella maniera più corretta il contributo dei CAT Tools all'interno della didattica. La tecnologia non è semplicemente una delle diverse categorie in cui un traduttore deve dimostrare capacità e padronanza. Un utilizzo attivo e intelligente di TM

e TA produce inevitabilmente delle modifiche nella natura e nel bilanciamento di tutte le altre componenti, quindi in ultima istanza nel profilo professionale del traduttore (Pym, 2013: 492).

#### **4.2.2. Conoscenze tecniche e competenza traduttiva**

Le tecnologie devono quindi essere considerate come una parte a sé stante della composita competenza traduttiva? A complicare ulteriormente la discussione contribuisce la loro natura in continua evoluzione, che fa sì che essi richiedano un continuo aggiornamento, per cui già al momento dell'acquisizione del loro utilizzo saranno state sviluppate tecnologie più avanzate che faranno diventare la conoscenza acquisita già obsoleta. La vera competenza relativa ai *tools* per la traduzione non risiederà quindi nell'imparare ad utilizzarne uno in particolare, quanto nello sperimentare diversi strumenti ed essere predisposti ad un rapido apprendimento di versioni sempre nuove tramite lo sviluppo di una percezione intuitiva o del ricorso a supporti specifici come gruppi di supporto online, tutorial o forum di discussione. Non si tratta quindi di imparare ad usare uno strumento, ma di apprendere una metodologia di lavoro che tenga in considerazione l'apporto delle risorse digitali e le sappia sfruttare al meglio, con una costante apertura al nuovo. Pym (2003: 491) propone una definizione minimalista del concetto di competenza traduttiva. Secondo l'autore, essa risiede nell'abilità di generare una serie di possibili frasi in una lingua di arrivo corrispondenti ad una frase nella lingua di partenza, e nell'abilità di selezionare solo la frase più appropriata dalla serie generata, in maniera rapida e con un elevato grado di sicurezza. Insieme, queste due abilità rappresentano per l'autore il cuore della competenza in traduzione, perché riferite specificamente al processo traduttivo e a nient'altro, pur non escludendo la lunga e diversificata lista di conoscenze che il traduttore deve possedere per svolgere al meglio il proprio lavoro (tra le quali, le conoscenze tecnologiche specifiche). Per Pym questo processo pragmatico di generazione e selezione è il fulcro di tutto, un meccanismo di *problem-solving* che spesso avviene in maniera apparentemente automatica. L'approccio minimalista aiuta a chiarire la posizione dei CAT Tools e delle altre tecnologie per la traduzione, ricordando che esse non sono il fine della formazione, ma semplicemente strumenti al servizio delle due attività fondamentali del processo traduttivo, quella di generazione e quella di selezione. Gli strumenti elettronici rendono più rapida e più agevole la prima, estraendo in maniera automatica alternative dalle memorie di

traduzione, dai glossari e dizionari di riferimento e tramite il processo di TA, mentre favoriscono la seconda rendendo più facile l'identificazione dell'alternativa più adatta tramite diversi criteri di selezione come i contesti di utilizzo o il conteggio statistico delle occorrenze. Essi in sostanza non alterano il processo traduttivo, che rimane basato su attività di produzione e selezione, ma contribuiscono ad esso nel tentativo di semplificarlo e velocizzarlo. L'approccio minimalista di Pym pone l'accento sulla necessità di investire in particolare sull'abilità di selezione (capacità fondamentale e necessariamente umana) data l'abbondanza e la facilità di recupero di alternative dai supporti multimediali, e la tendenza affermata con l'utilizzo delle TM di adottare gli *exact match* in maniera indiscriminata, ovvero senza la fase valutativa che porta alla selezione della migliore soluzione. La didattica dei CAT Tools deve quindi essere pedagogicamente incentrata sullo sviluppo delle capacità umane di *problem-solving* e *decision-making*, deve insegnare che la parola finale spetta al traduttore e che gli strumenti devono essere utilizzati come tali.

#### **4.2.3. Alcuni studi ed indagini**

Come già visto, la scelta di inserire corsi sui CAT Tools all'interno dei curricoli universitari è principalmente dettata da un'istanza di formazione professionale che possa consentire agli studenti di affacciarsi in maniera preparata al mondo del lavoro (Bolaños - Medina e Gonzales, 2012:276). Da diversi studi sulle attitudini degli studenti di traduzione nei confronti dell'offerta formativa in generale e dei corsi sulle tecnologie applicate alla traduzione in particolare (Kenny e Way, 2001; Li, 2002; Bolaños - Medina e Gonzales, 2012) è emerso che gli studenti reputano in genere la propria formazione accademica insufficiente per il loro avviamento professionale, lamentando una scarsa esperienza pratica o una inadeguata preparazione agli aspetti più pragmatici della professione, ovvero quelli che prescindono dalla capacità di sapere tradurre. Come riconosciuto da Li (2002) che ha condotto uno studio per verificare quanto i corsi di traduzione cinese – inglese tenuti presso l'università di Honk Kong corrispondessero alle reali aspettative degli studenti e quali fossero le motivazioni che spingevano gli studenti a scegliere il curriculum in traduzione, si notano in genere due tendenze. Da un lato gli studenti che intendono in futuro lavorare nel settore della traduzione, dall'altro quelli mossi essenzialmente dal desiderio di padroneggiare al meglio due lingue e che considerano i curricoli in traduzione più adatti a tal fine per la loro natura più pratica.

Un'opinione comune è che attraverso la traduzione si possa fare esperienza degli utilizzi reali delle lingue, per cui anche chi non intende diventare traduttore considera questi corsi più adatti a possibili sbocchi occupazionali. Inoltre, all'interno nel programma di studi, le materie considerate più *utili* (come conoscenze spendibili concretamente) sono state quelle di natura più pratica, come la traduzione tecnico – scientifica o i laboratori di interpretariato, che consentono di fare esperienza diretta del processo traduttivo. Al contrario, i corsi teorici come la storia o la teoria della traduzione hanno ottenuto una valutazione piuttosto negativa in termini di utilità intesa come spendibilità pratica per uno sbocco professionale. In generale, gli studenti hanno affermato di non trovare che il programma di studi rispecchiasse le esigenze del mercato in maniera soddisfacente (i due terzi dei partecipanti); l'interesse dei partecipanti all'indagine è principalmente quello di poter concentrare i propri sforzi nello studio e nella formazione su quelle nozioni e quelle competenze che potranno rivelarsi vincenti al fine di inserirsi nel mercato del lavoro.

Un altro interessante questionario è quello condotto da Bolaños - Medina e Gonzales (2012), che mira a comprendere l'attitudine degli studenti circa i CAT Tools al fine di individuare gli aspetti sui quali insistere per una migliore didattica di tali strumenti, anticipare le difficoltà di apprendimento e identificare le strategie che favoriscano migliori attitudini di apprendimento. In generale gli studenti riconoscono che questi strumenti accelerano il lavoro e ne migliorano alcuni parametri qualitativi, ma viene anche riconosciuto il bisogno di possedere un'adeguata dimestichezza con l'utilizzo del computer e delle tecnologie, che diventa il fattore di discriminazione tra coloro che si relazionano con questi strumenti in maniera serena e quelli che invece ne sono intimoriti proprio per la loro scarsa conoscenza informatica. Questi ultimi percepiscono l'utilizzo dei *tool* come fonte di stress, e sviluppano nei loro confronti un'attitudine negativa. Anche la varietà degli strumenti e la loro continua evoluzione viene vista come un ostacolo, fino al punto da far desistere un 10% circa degli intervistati dall'idea di farne uso.

Per concludere, i CAT Tools e le tecnologie applicate alla traduzione sono ormai tra i requisiti indispensabili per accedere al mercato della traduzione, pertanto essi vanno ad aggiungersi alla lunga lista di abilità e conoscenze che gli studenti in traduzione dovranno acquisire durante la loro formazione. Affinché la loro conoscenza contribuisca alla creazione di una competenza traduttiva, occorre che gli studenti ricordino di considerare le tecnologie come strumenti, da cui trarre il massimo del supporto ma di cui

fidarsi in maniera oculata, e che siano consapevoli di come la responsabilità dell'intero processo rimanga comunque del traduttore. Una didattica consapevole delle implicazioni cognitive discusse nel Terzo capitolo di questa tesi consentirà agli studenti di sviluppare un corretto approccio a tali strumenti, limitando le eccessive influenze come nel caso dell'adozione indiscriminata dei suggerimenti ottenuti, riconoscendo i contributi positivi che essi possono offrire e fornendo sufficienti conoscenze che possano mettere in guardia contro gli ostacoli che essi pongono alle volte al naturale processo mentale di traduzione.

### 4.3 *CAT Tools strumenti di analisi linguistica*

#### 4.3.1. **Un confronto con i corpora**

Come più volte affermato, i CAT Tools vengono in genere considerati come strumenti di uso pratico e mirato, la cui utilità è fondamentalmente economica ovvero diretta al risparmio di tempo, lavoro e denaro. Sebbene venga anche loro riconosciuto di contribuire positivamente alla qualità della traduzione in termini di accuratezza e coerenza terminologica<sup>25</sup>, molto spesso viene trascurato un aspetto importante e particolarmente significativo soprattutto in ambito didattico e di ricerca. Si tratta della possibilità di utilizzare tali strumenti come mezzo di analisi e osservazione linguistica a partire dal bagaglio di risorse contenute nelle memorie di traduzione, fruite e analizzate secondo il criterio su cui si basano i modelli e le tecniche di osservazione tipici della linguistica dei corpora. I corpora sono ampi database di testi in *machine-readable format* che costituiscono degli esempi reali e rappresentativi di utilizzo della lingua, raccolti in base ad una precisa serie di criteri tematici o linguistici. Un corpus contiene centinaia di milioni di parole al suo interno, ed è proprio la sua ampiezza e il continuo arricchimento dei suoi contenuti che lo rendono altamente rappresentativo della lingua o delle lingue che esso include (Sharoff, 2004: 23). Attraverso un complesso processo di estrazione e ricerca terminologica, il corpus è in grado di fornire informazioni diverse sui comportamenti linguistici desunti dai testi che esso raccoglie. Tali informazioni potranno essere utili al fine di condurre ricerche linguistiche relative a una lingua o a precise tipologie testuali o ambiti linguistici settoriali, in particolare relativamente alle

---

<sup>25</sup> Il miglioramento qualitativo della traduzione è tra l'altro un parametro discutibile perché, come visto nel capitolo precedente, la rigida suddivisione in frasi compiuta dai CAT Tools può inficiare la naturalezza del testo tradotto e in ultima istanza la qualità dello stesso.

caratteristiche terminologiche e sintattiche, ai costrutti più frequenti, alla presenza di collocazioni o espressioni specifiche etc. (Gandin, 2009: 133). Esistono come è noto corpora monolingui, bilingui o plurilingui. Questi ultimi due sono nella maggior parte dei casi anche paralleli, ovvero contenenti dei testi originali in una determinata lingua, e la loro relativa traduzione in una o più lingue, allineati frase per frase secondo la stessa logica descritta per le TM. Inoltre, i corpora paralleli possono essere uni-direzionali come appena descritto, o bi-direzionali, laddove i testi originali non siano solo in una delle lingue contenute, ma in più di una (es. testi originali inglesi allineati alla loro traduzione italiana e testi originali italiani allineati alla loro traduzione inglese). Corpora di soli testi originali o di sole traduzioni vengono invece definiti comparabili. Di fatto, molti degli strumenti offerti dai CAT Tools sono stati sviluppati proprio a partire dalle funzioni fondamentali dei *corpus analysis tools*, gli appositi software utilizzati per interrogare il corpus e visualizzare i risultati estratti. In particolare, sono proprio i corpora bilingui o paralleli a condividere molte caratteristiche con le *translation memories* (Zanettin, 2002:2 e Ciola, 2008:1; Bowker e Barlow, 2004). È possibile quindi proporre in ambito didattico un confronto tra i due strumenti, al fine di comprendere i criteri secondo i quali sono stati sviluppati i primi (quindi anche le loro potenziali applicazioni), a partire dalla riflessione teorica e pratica sviluppata attorno ai secondi. Dal punto di vista formale, entrambi i sistemi dividono i testi in segmenti allineati con le proprie traduzioni. La capacità comune ad entrambi i sistemi di scomporre i testi in porzioni minime, denominata *chunking*, rende possibile una ricerca minuziosa degli elementi lessicali e sintattici e uno studio approfondito delle caratteristiche testuali. Essa funziona come una sorta di "lente di ingrandimento" che permette di mettere a fuoco porzioni più o meno ampie di testo e di dedicare quindi un'attenzione particolare a tutti i livelli, da quello terminologico a quello della costruzione di frasi e periodi (Gandin, 2009:141). Nei testi bilingui, l'applicazione del chunking può agevolare il confronto tra le due lingue, tra due modi di codificare i concetti e tra le caratteristiche del linguaggio settoriale preso in esame. La divisione in chunk suggerisce inoltre una riflessione sul concetto stesso. Secondo la linguistica cognitiva, il chunk è l'unità minima di informazione assimilabile dalla mente umana. Per la linguistica dei corpora, il chunk è la porzione minima di testo che consente di conoscere l'utilizzo reale di una parola in una determinata lingua, grazie alla sua presenza all'interno di un contesto (Calzolari e Lenci, 2004). Nei CAT Tools tali significati si fondono: la

memoria di traduzione immagazzina non parole, ma segmenti di testo e, al fine di comprenderne il significato, rileva non le singole parole, ma ripropone i segmenti interi come unità di senso e di discorso.

Pur riconoscendo le ovvie differenze, è possibile pensare alla TM come a un piccolo corpus bilingue, contenente testi allineati facenti riferimento ad un unico settore tecnico-scientifico. Così come i corpora bilingui, le TM consentono di osservare due lingue a confronto, cosa che può rivelarsi utile sia per lo studio e l'apprendimento della lingua straniera che per la formazione in traduzione. L'insieme dei segmenti allineati contenuti nella TM può essere infatti interrogato non solo al fine di ottenere un "suggerimento" durante la fase di traduzione ma - proprio come avviene di norma attraverso gli strumenti di analisi dei corpora - anche per avere accesso ad uno spaccato di lingua autentica, dal quale desumere l'uso reale di alcuni termini nel loro contesto e l'occorrenza di determinati *pattern* linguistici che caratterizzano un dato ambito specialistico. Grazie alla loro estensione, i corpora assicurano una sicura rappresentatività del linguaggio reale. Indubbiamente una TM non potrà eguagliare l'estensione di un corpus, pur essendo potenzialmente estendibile all'infinito. Ferma restando la necessità di utilizzare delle TM sufficientemente ampie da fornire risultati significativi e rappresentativi, tale aspetto viene ad essere compensato per certi versi da alcune caratteristiche proprie delle memorie di traduzione. Innanzi tutto, le TM offrono il vantaggio di essere nella maggior parte dei casi altamente settoriali, perché come si è visto è prassi comune creare e adoperare TM specifiche non solo per ogni settore tecnico - scientifico, ma alle volte persino per ogni cliente o per ogni prodotto o argomento specifico. È verosimilmente ipotizzabile, per esempio, che un traduttore medico non lavori con un'unica TM generica, ma che ne posseda diverse per ogni singola branca della medicina con la quale egli lavora più di frequente. Il testo contenuto all'interno della memoria di traduzione è quindi strettamente afferente ad un ambito ben determinato, e ciò rappresenterà il vero valore aggiunto delle TM. I corpora dal canto loro, proprio per la vastità di tipologie testuali e ambiti contenuti, non possono garantire un'altrettanta specificità dell'output ottenuto, ed esistono ad oggi pochissimi corpora paralleli dedicati ai linguaggi specialistici, come il TRANSEARCH o l'MLCC (Gandin, 2009: 137). Inoltre, ogni doppietta di segmenti allineati contenuta all'interno della memoria, prima di essere salvata subisce diverse fasi di revisione, editing e validazione finale. Si può quindi

affermare che l'output estratto dalle TM garantisca una elevata percentuale di affidabilità e correttezza che non necessita di essere confrontata e avallata da numeri molto elevati di occorrenze, come invece avviene per le ricerche svolte tramite *corpus analysis tool*, che implicano un calcolo statistico e la verifica di sufficienti occorrenze per determinare la reale usabilità di una data espressione.

Al di là dei vantaggi dei singoli strumenti, ci interessa qui suggerire che l'introduzione di un confronto tra corpora e CAT Tools nella formazione universitaria potrebbe ampliare gli orizzonti dello studio circa i così detti "sussidi per la traduzione", offrendo ai discenti uno strumento per la creazione di un modello di analisi lessicale, sintattica e testuale grazie ad una applicazione rinnovata e riadattata dei modelli di analisi della linguistica computazionale e dei corpora<sup>26</sup>, che sfrutti al meglio le potenzialità dei CAT Tools. Le informazioni ottenute da corpora e TM sono per ovvi motivi diverse, e non si intende qui equiparare gli strumenti quanto piuttosto tracciare un filo conduttore tra i criteri che regolano l'utilizzo di entrambi. In tal modo, se sufficientemente preparati all'interno dell'ambiente didattico, gli studenti potranno utilizzare il loro software di traduzione assistita per attuare ricerche simili a quelle che potrebbero compiere su di un corpus (Sassi e Ceccotti, 2000), ottenendo informazioni utili tanto per la loro formazione linguistica quanto per lo sviluppo di una competenza traduttiva.

#### **4.3.2. Funzioni di ricerca**

Per dare man forte alla proposta appena esposta, verranno adesso esaminate alcune delle funzioni maggiormente impiegate nei sistemi di analisi dei corpora, e confrontate con delle applicazioni simili implementate nei software di traduzione assistita. Rendere esplicita tale similarità in ambito didattico può consentire un approccio più consapevole alle reali potenzialità fornite da queste ultime applicazioni.

LISTE DI FREQUENZA. Sono elenchi di parole che vengono estratte insieme al conteggio della loro frequenza, ovvero il numero di occorrenze (le volte in cui esse compaiono) all'interno di un corpus di testi. Le *frequency list* consentono inoltre di calcolare il rapporto tra i *token* – il numero totale delle parole del corpus – e *type* - cioè il numero delle attestazioni dell'utilizzo dei singoli token all'interno del corpus (Bowker

---

<sup>26</sup> La linguistica computazionale offre un approccio probabilistico e statistico all'analisi testuale, mentre la linguistica dei corpora rende possibile uno studio basato su dati empirici e regolari. I due metodi sono strettamente correlati, come si deduce da Chiari, 2008: 76.

2002: 155), come visto in 1.4. Il calcolo percentuale tra il numero di type e il numero di token può fornire informazioni, ad esempio, sulla “densità” testuale oppure sulla lunghezza media delle parole, contribuendo all’individuazione delle caratteristiche lessicali dei testi presi in esame. Inoltre, è possibile considerare le forme flesse, ad esempio i plurali dei nomi, come dei token separati oppure ricondurli alla forma base e recuperarli insieme agli altri type, in base al tipo di informazione che si intende ottenere. I dati relativi alla frequenza terminologica permettono anche di individuare i termini aventi maggiore rilevanza all’interno di un corpus, potendo così identificare anche i concetti fondamentali relativi a una disciplina. Parallelamente, alcuni strumenti di estrazione terminologica messi a disposizione dei CAT Tools, come per esempio SDL Multiterm Extract<sup>27</sup>, seguono lo stesso modello di analisi, ricercando all’interno dei testi i termini più rilevanti in termini di frequenza, permettendo di catalogarli in un glossario. Il traduttore può quindi sfruttare l’applicazione basilare di tale software per ottenere un bilancio delle parole chiave di un relativo ambito e il loro inserimento in un database terminologico. È tuttavia possibile ipotizzare un ampliamento di tale utilizzo, già di per sé importante, in ambito didattico, al fine di fornire agli studenti un potente strumento di ricerca, che sia capace di sviluppare la riflessione linguistica e terminologica, grazie al recupero di un’ampia varietà di informazioni, contenuti e definizioni, allo stesso modo in cui avviene nel caso dei corpora (Chiari, 2008:70). In questo modo, essi potranno acquisire un metodo di studio ed analisi, utile a formare una solida conoscenza dei linguaggi specialistici, a comprenderne le strutture tipiche e a produrre un bagaglio di informazioni che sarebbe altrimenti più difficilmente individuabile, in quanto richiederebbe molto più tempo e maggiori sforzi.

PROGRAMMI PER LE CONCORDANZE. Si tratta di strumenti che permettono la ricerca, all’interno di una raccolta di testi, di tutte le occorrenze di una particolare parola. Essi sono particolarmente utili perché permettono di determinare le modalità di impiego di un certo termine, tramite la sua osservazione in uno o più contesti di utilizzo messi a confronto. Ciò si rivela utile tanto per i linguaggi specialistici quanto per il linguaggio di uso comune, che essendo più ricco di termini polisemici ha maggiore bisogno di informazioni contestuali per disambiguare (Sharoff, 2007: 16). Nei sistemi di analisi dei corpora, il formato più comune di visualizzazione è definito KWIC (*Key Word In*

---

<sup>27</sup> <http://www.sdl.com/products/sdl-multiterm/extract.html>

*Context*), in cui tutte le occorrenze della parola ricercata sono allineate al centro della finestra. Anche i software per la traduzione assistita sfruttano i vantaggi di tale funzione, proponendo il *Concordance* (cf. 2.1), uno strumento molto importante di ricerca ed utilizzo coerente della terminologia all'interno dei segmenti salvati nella memoria di traduzione. La ricerca può essere compiuta in maniera istantanea, sia su termini che su espressioni più ampie, appartenenti al testo source o a quello target. Durante la traduzione, il beneficio principale di questo strumento è quello di fornire una risposta immediata a eventuali dubbi del traduttore, soprattutto nel caso in cui serva disambiguare tra le diverse accezioni di un termine polisemico osservandone l'utilizzo all'interno del contesto pertinente al testo che si sta traducendo. A differenza dei sistemi KWIC, il *Concordance* non fornisce risultati incolonnati, ma evidenzia la parola ricercata, visualizzandola all'interno del segmento di appartenenza. Inoltre, poiché i testi vengono allineati sempre a livello di frase, occorre specificare che laddove si ricerchi una concordanza in un corpus bilingue o in una TM, il programma recupererà i segmenti di testo in cui compare la parola ricercata ed i rispettivi segmenti nella lingua della traduzione, ma non sarà in grado di identificare all'interno di essi la parola corrispondente a quella ricercata<sup>28</sup>. Questa dovrà essere desunta dal traduttore sulla base delle sue conoscenze personali. In ogni caso, l'aver ben compreso il contesto di utilizzo di un termine renderà più semplice l'identificazione della corretta traduzione. Come si è detto, le TM sono generalmente settoriali, quindi il problema della disambiguazione tra più contesti viene a mancare. Tuttavia, alcuni termini assumono un significato diverso all'interno di ambiti specifici, e necessitano comunque di essere ben contestualizzati. L'osservazione del loro utilizzo all'interno di frasi esemplificative è da questo punto di vista molto prezioso.

FUNZIONE DI RICERCA DELLE COLLOCAZIONI. È strettamente legata alla funzione precedente, perché si avvale di essa per ricercare quelle parole che in un corpus di testi possono, con alta probabilità, trovarsi spesso insieme, costituendo un'unità lessicale. All'interno dei *corpus analysis tools*, l'analisi delle collocazioni consente di recuperare esempi di strutture sintattiche codificate, come ad esempio verbo-preposizione, verbo-sostantivo che non si incontrano in maniera casuale all'interno di più testi, ma sono standardizzate per esprimere un concetto ben preciso. Tale funzione ha, inoltre, il compito

---

<sup>28</sup> Il così detto *term matching* (Ciola 2008: 6) può essere svolto da sistemi specifici che ricorrono a calcoli statistici per individuare la parola che ricorre più di frequente nelle traduzioni dei segmenti che contengono il termine ricercato.

di individuare i costrutti sintattici tipici dei linguaggi specialistici, ovvero le *multi-word units*, gruppi di due o più parole utilizzate per esprimere, secondo una logica compositiva, un unico referente (Sharoff, 2007:16). È anche possibile ricercare le possibili co-occorrenze, definite come il calcolo delle probabilità che hanno le parole di un testo di essere associate, all'interno del testo stesso, ad una o più parole prestabilite (Sassi e Ceccotti, 2000). Il traduttore che lavora in un particolare ambito tecnico dovrà essere in grado di riconoscere e comprendere tanto le unità polisemiche quanto le possibili co-occorrenze, per garantire una corretta interpretazione e identificazione dell'appropriato equivalente traduttivo. Esprimendo dei concetti monoreferenziali, tali termini possiedono la maggior parte delle volte un unico traduttore, che ogni traduttore deve utilizzare per garantire la massima precisione, pena una comunicazione non efficace. È auspicabile, quindi, che un curriculum didattico attento alla formazione linguistica educi gli studenti al fine di far comprendere loro l'importanza delle collocazioni all'interno di un testo specialistico. La loro individuazione consente un approccio più consapevole e professionale al testo. A tal fine, è possibile sfruttare le potenzialità di Multiterm, il quale viene normalmente utilizzato al solo scopo di creare dei glossari le cui voci sono costituite da singole parole e i loro equivalenti nelle lingue di arrivo. Tuttavia, tale software può essere considerato più di un semplice strumento terminologico, un database contenente *chunks* lessicali caratterizzanti i linguaggi settoriali e fondamentali per un approccio competente alla traduzione (Chang *et al.*, 2006:3). Se da un lato la linguistica computazionale è in grado di estrapolare dai corpora le collocazioni, al fine di analizzare l'utilizzo ricorsivo di determinate unità lessicali nei testi settoriali, dall'altro Multiterm (e i software equivalenti forniti da altri CAT Tools) permetterebbe di guardare a tali strutture in maniera sistematica e codificata, riconoscendole quali unità semantiche complesse. Così come nella linguistica computazionale è possibile interrogare i corpora al fine di recuperare le collocazioni, allo stesso modo i CAT Tools, attraverso la funzione denominata *Active Term Recognition*, potrebbero svolgere una ricerca simile. Utilizzando la TM come un corpus, e il glossario come una raccolta di sintagmi nominali, essi sono in grado di attuare l'analisi di un segmento sorgente al fine di riscontrare, sulla base dei dati salvati in memoria, tali costrutti lessicali, che sono riconosciuti come un'unica espressione alla quale è associata la corrispondente traduzione nella lingua target.

Alla luce di quanto detto, è evidente come la riflessione linguistica sia un tassello fondamentale della formazione degli studenti di traduzione. Essa permette di guardare alla traduzione non come un'attività meccanica o fine a se stessa, bensì come un'occasione di apprendimento, di accrescimento delle proprie competenze, di analisi e studio del linguaggio, delle sue evoluzioni ed applicazioni.

#### 4.4 *A sostegno dell'apprendimento dei linguaggi specialistici*

All'interno di un ambiente didattico, le tecnologie possono essere acquisite in maniera fine a sé stessa, come in un corso pratico di CAT Tools, oppure possono essere adoperate al fine di acquisire attraverso di esse altre competenze. Forse quest'ultima è la modalità di apprendimento più interessante, perché presenta agli studenti una immediata utilità dello strumento che stanno imparando ad usare, ne consente un utilizzo reale e non simulato e spinge gli studenti ad adoperare un approccio ragionato e critico nei suoi confronti (Bowker, 2002: 17).

##### 4.4.1. *I Languages for Special Purposes*

Come già emerso nel corso della trattazione al precedente paragrafo, i linguaggi specialistici sono una componente molto importante della formazione in traduzione. Anche noti come *LSP – Languages for Special Purposes*, essi vengono definiti come lingue utilizzate “*to discuss specialized fields of knowledge*” (Bowker e Pearson, 2002: 25). Diversamente dal linguaggio corrente (*LGP – Language for Generic Purposes*), che viene invece utilizzato “*to talk about ordinary things in a variety of common situations*” (*ibid.*), i linguaggi specialistici possono includere temi che riguardano qualsiasi attività, professionale e non, che sia caratterizzata da un elevato grado di specificità dovuto alla presenza di un vocabolario specializzato e utilizzato solo in tale settore, e di collocazioni e aspetti stilistici e grammaticali tipici ed esclusivi di tale contesto specialistico (Gandin, 2009: 137). La capacità di saper identificare le caratteristiche specifiche di un linguaggio specialistico nella lingua di origine e di riportarle in maniera equivalente nella lingua di arrivo è fondamentale per un traduttore, e deve essere allenata già durante la sua formazione. L'elevato grado di specializzazione che richiede oggi il mercato della traduzione è uno dei parametri professionali ai quali gli studenti devono poter puntare in maniera consapevole. Sebbene l'esperienza sia la vera maestra nella maggior parte dei

casi, poter affrontare i primi più difficili anni di lavoro potendo contare su di una buona preparazione è sicuramente un vantaggio. Tale preparazione implica anche che, in una situazione ideale, gli studenti possano maturare già durante gli studi una loro personale predilezione per uno specifico settore, avendo avuto modo di venire a contatto e di confrontarsi in maniera pratica con diversi linguaggi specialistici già durante il periodo di formazione. I traduttori altamente specializzati sono molto apprezzati dalle agenzie di traduzione, che richiedono infatti in molti casi di specificare, al momento della candidatura di ogni nuovo traduttore freelance, i principali campi di esperienza. I maggiori atenei in cui si studia traduzione offrono corsi di traduzione tecnico – scientifica, generalmente di ambito medico, legale, economico e ingegneristico<sup>29</sup>, e non di rado questi corsi così mirati si avvalgono delle tecnologie CAT per permettere agli studenti di sperimentare le reali dinamiche di lavoro e ancor più di cominciare a raccogliere materiali utili, come i primi glossari e le prime TM, prodotti e utilizzati durante le esercitazioni. Inoltre, le numerose funzioni di utilizzo, verifica e manutenzione della terminologia vengono in soccorso nella gestione di tale importante aspetto del lavoro di traduzione, rappresentato dalla scelta e dall'utilizzo omogeneo della terminologia e delle espressioni esatte quali varianti lessicali, termini composti e costrutti sintattici specifici dell'ambito per il quale si traduce. Va anche tenuto in considerazione che la manutenzione e l'aggiornamento costante della propria terminologia e delle memorie è fondamentale data la rapida e continua evoluzione dei linguaggi specialistici, della qual cosa gli studenti dovranno essere resi coscienti e capaci, fornendo loro padronanza di quegli strumenti che li aiuteranno a curare tale aspetto (Tufis *et al.*, 2004: 163).

#### **4.4.2. Lo studio tramite i corpora e i CAT Tools**

La preoccupazione dominante dei docenti nell'apprendimento dei linguaggi specialistici riguarda quasi sempre l'apprendimento della terminologia (Maia, 2005), essendo proprio il lessico l'aspetto più particolare dal quale emergono le differenze più evidenti rispetto al linguaggio comune (Mazzotta, 2007: 138). Tuttavia, per una comprensione piena dei linguaggi specialistici non sono tanto i termini a essere

---

<sup>29</sup> L'Alma Mater Studiorum di Bologna e la Scuola Superiore per interpreti e traduttori di Trieste, i due centri principali in Italia per la formazione dei traduttori, offrono entrambe un intero corso di laurea magistrale dedicato alla traduzione specializzata (<http://corsi.unibo.it/Magistrale/TraduzioneSpecializzata/piano-di-studio-anno-accademico-20142015.htm>; <http://www.sslmit.units.it/images/page/229/piano%20di%20studi%20LM94%20esteso.pdf>).

importanti, quanto la comprensione di una serie di procedure morfo-sintattiche frequenti nel discorso scientifico in maniera preponderante rispetto alla lingua comune. Lo stile nominale, le costruzioni passive e impersonali, le frasi implicite gerundive, infinitive e participiali sono alcuni degli elementi più caratteristici (*ibid*), che occorre imparare a riconoscere, e riprodurre correttamente nella lingua verso cui si traduce (in genere la propria lingua madre). La linguistica dei corpora offre come si è visto notevoli vantaggi per la comprensione, l'apprendimento e l'analisi dei linguaggi specialistici, grazie all'osservazione di testi autentici che forniscono uno spaccato reale della lingua in uso corrente in un determinato settore. Tramite i corpora è possibile individuare una terminologia specialistica accurata attraverso la generazione di liste di parole presenti nel corpus e la relativa frequenza, potendo identificare quelle più ricorrenti; è possibile identificare le collocazioni terminologiche tipiche del genere analizzato e studiare le strutture grammaticali, sintattiche e stilistiche proprie di un linguaggio specialistico. I corpora paralleli consentono inoltre di poter estrarre terminologia bilingue e di osservare costrutti e sintassi tipici dello stile dei linguaggi analizzati nelle due lingue. Tuttavia, l'applicazione della *corpus analysis* allo studio dei testi tradotti è un fenomeno di ricerca recente della linguistica dei corpora e ancora poco esplorato, soprattutto per quanto riguarda i linguaggi specialistici (Gandin 2009: 134), data la limitata disponibilità, come si è visto, di corpora settoriali sufficientemente ampi e rappresentativi. I CAT Tools, sviluppati dai corpora ed espressamente destinati ai linguaggi di tipo settoriale e specialistico, sono validi strumenti per la formazione degli studenti circa i linguaggi specialistici. Essi offrono un triplice vantaggio: rendono consapevoli gli studenti dell'importanza dei LSP; consentono la raccolta, l'utilizzo e la gestione di risorse linguistiche preziose e possono essere utilizzati al fine di analizzare, studiare e confrontare le caratteristiche morfosintattiche e terminologiche dei linguaggi specialistici nelle due lingue contenute in una TM.

Il primo aspetto appena menzionato – l'importanza dei linguaggi specialistici – sarà evidente già dall'impostazione della didattica: laddove infatti si adottino tecnologie CAT non si potranno che utilizzare testi di carattere tecnico – scientifico, più adatti ad essere elaborati tramite tali strumenti. Lavorando con testi specialistici gli studenti ne identificheranno le caratteristiche fondamentali che i CAT Tools rendono ancor più evidenti: il relativo tasso di ripetitività, l'abbondanza di espressioni standardizzate,

l'utilizzo di una terminologia puntuale. La pratica reiterata e l'esperienza diretta renderanno esplicite anche le caratteristiche morfo-sintattiche, lo stile e le tipologie testuali dalle caratteristiche espressive ben precise.

Relativamente al secondo aspetto, un esempio di come i CAT Tool possano venire in aiuto è suggerito da Nagata *et al.* (2001), che parlano della possibilità di estrarre terminologia bilingue direttamente dal web, in particolare da quelle pagine di settori specialistici in cui i termini settoriali nella lingua della pagina sono spesso accompagnati dai loro equivalenti in inglese. La metodologia proposta dagli autori prevede la ricerca congiunta di termini estrapolati da dizionari settoriali e della loro traduzione, tramite motore di ricerca sul web. I risultati ottenuti, analizzati e valutati per accertarne l'affidabilità, possono essere salvati all'interno di glossari creati con gli appositi applicativi della suite, come Multiterm, e corredati da informazioni sulla fonte e sul contesto di utilizzo riscontrato. È anche possibile ricavare dal web interi testi bilingui, allinearli e salvarli nella TM. Una volta sufficientemente popolata, questa potrà essere utilizzata tanto per tradurre, quanto per svolgere delle ricerche al fine di trarre informazioni sulle caratteristiche del settore preso in considerazione. Sempre utilizzando il web come un enorme corpus è possibile pensare a delle modalità di verifica della terminologia o dei costrutti morfosintattici tramite un confronto tra la terminologia estratta dalla TM o dai glossari del traduttore con quella reperibile in internet all'interno dei corpora online o di pagine web attendibili (Ciola, 2008: 10). Il risultato potrebbe essere un report che elencherà i termini analizzati e la loro ricorrenza (zero, *n*) all'interno delle risorse consultate.

Infine, il terzo aspetto che si evince già in parte da alcune delle considerazioni appena espresse, implica il confronto tra le due lingue. Sia le TM che i sistemi di gestione della terminologia possono essere molto utili a tal fine. Le esercitazioni che possono essere proposte agli studenti sono mirate a far emergere analogie e differenze tra i linguaggi specialistici nelle due lingue e ancor prima le caratteristiche peculiari di ciascuna di esse. Ciò sarà possibile attraverso ricerche compiute tramite la funzione *Concordance*, per osservare i diversi livelli linguistici, dalla composizione della frase alla sintassi fino al livello delle parole. Si potranno annotare dei glossari con informazioni contestuali o formali utili al momento di tradurre, come esempi di uso di un dato termine o espressione. Si potrà inoltre estrarre terminologia tramite l'apposito applicativo anche a scopo di auto-

valutazione, per verificare che le scelte terminologiche e stilistiche compiute durante una traduzione con l'ausilio della TM siano corrette. In generale, la pratica e la gestione dei *tool* in maniera personale e finalizzata a uno scopo ben preciso non solo contribuiscono alla formazione di una procedura operativa utile per la gestione del proprio lavoro, ma agevolano anche l'apprendimento linguistico attraverso l'iterazione e la pratica. Infine, dal confronto tra le due lingue potranno essere identificate le più idonee strategie traduttive, inizialmente studiate proprio dalla stilistica comparativa (Wotjac 1997: 102). La scelta delle strategie traduttive è legata al controverso concetto di equivalenza, che si declina su livelli diversi, dall'equivalenza lessicale e sintagmatica – semantica al livello più ampio dell'equivalenza comunicativa (*ibid.*). Avendo ogni linguaggio specialistico in ogni lingua le sue specifiche caratteristiche, la ricerca di equivalenze a qualsivoglia livello implica, oltre ad una solida conoscenza delle lingue e dei linguaggi settoriali, anche la capacità di saper applicare le giuste scelte e strategie, che garantiscano una equivalenza di tipo funzionale data la natura essenzialmente pragmatica della maggior parte dei testi tecnici – scientifici.

#### 4.5 *CAT Tools all'Università degli Studi di Palermo*

##### **4.5.1. La prevalenza dell'approccio pratico**

L'ingresso a pieno titolo delle tecnologie per usi specifici all'interno dei curricula universitari in lingue e traduzione non è più recentissimo. Soprattutto negli ultimi dieci anni i laboratori pratici di CAT Tools hanno conquistato il loro spazio presso i corsi a indirizzo linguistico delle maggiori università italiane. Tuttavia, anche nei centri che hanno dedicato più spazio alle dinamiche della *Computer-Aided Translation*, l'approccio prevalente è sempre quello pratico e funzionale, e poco spazio viene concesso alla sperimentazione proposta nei capitoli e paragrafi precedenti. Gli strumenti di sussidio alla traduzione vengono visti nella maggior parte dei casi come importanti e fondamentali strumenti, ma nulla di più. A testimonianza di ciò, la scarsità di letteratura a loro riguardo prodotta in Italia, a fronte di un maggiore interesse dimostrato in altri paesi europei, americani ed asiatici. Ciò può essere in parte dovuto al permanere di approcci più classici e teorici alla traduzione, in parte anche al prevalere della funzionalità pratica degli strumenti digitali, ma quello che forse determina maggiormente questo stato di cose è la

mancata armonizzazione dell'aspetto pratico e di quello teorico che fa desistere gli studiosi dal considerare i CAT Tools interessanti anche come oggetti di studio e come strumenti per una didattica linguistica e più strettamente traduttiva innovativa e sperimentale nella doppia accezione di nuova e pratica.

L'Università degli Studi di Palermo ha dedicato grande interesse ai CAT Tools negli anni passati, con l'introduzione di corsi laboratoriali all'interno degli indirizzi di laurea triennale e specialistici di lingua e traduzione<sup>30</sup>. Questi corsi, divisi in parti uguali tra lezioni frontali ed esercitazioni in laboratorio, hanno permesso agli studenti di inquadrare la realtà della traduzione professionale e, richiamando quanto appreso durante i corsi di lingua e traduzione circa i linguaggi specialistici, gli ambiti di applicazione delle tecnologie CAT e le tipologie testuali che più beneficiano del loro utilizzo. La parte pratica, svolta nei primi anni utilizzando Wordfast e successivamente SDL Trados, ha fornito le basi per l'utilizzo degli strumenti CAT, soffermandosi principalmente sui criteri e le finalità in base ai quali e in vista di cui si sviluppano le funzioni principali, al fine di consentire agli studenti di applicare le conoscenze acquisite anche ad altri software dello stesso tipo. La prova finale consisteva in una semplice traduzione svolta con l'ausilio del CAT Tool – volta all'osservazione dell'utilizzo di quest'ultimo, non tanto della correttezza della traduzione stessa – e in una discussione orale volta a testare il grado di comprensione delle conoscenze fornite nel corso delle lezioni. In linea generale, questa è l'organizzazione della maggior parte dei corsi di CAT Tools. L'attenzione all'aspetto linguistico è limitato al minimo, a meno che essi non siano inseriti all'interno di insegnamenti di lingua e traduzione come laboratori per le esercitazioni pratiche, o come si è visto, all'interno di corsi sui linguaggi specialistici.

#### **4.5.2. L'esperienza della Summer School**

Un approccio più interessante è stato sviluppato all'interno dell'ateneo palermitano negli anni a seguire. In particolare, dal 2008 al 2010 sono state organizzate tre edizioni di una Summer School internazionale intitolata "*CAT Tools: Translation, Localization, Internationalization*", della durata di due settimane e incentrata come dice il titolo sulla presentazione del settore della traduzione e della localizzazione, nonché sulle principali tecnologie di sussidio alla traduzione. Hanno preso parte alla Summer

---

<sup>30</sup> Si trattava in particolare dei corsi di laurea triennale in Lingue per il Web e in Traduzione, Italiano L2 e Interculturalità e del corso di laurea specialistico in Tecnologie e Didattica delle Lingue.

School, sia studenti italiani e palermitani, che studenti provenienti da alcuni atenei partner di Romania, Regno Unito e Austria<sup>31</sup>. Tra gli esperti invitati a presentare i diversi aspetti della “industria” della traduzione si sono avvicendati docenti universitari, traduttori professionisti, esperti in localizzazione e tecnici provenienti da diverse realtà di sviluppo software e strumenti digitali. Gli incontri con questi professionisti si sono svolti a mo’ di dibattito, con la possibilità da parte degli studenti di porgere domande ed esprimere osservazioni. La parte pratica del corso è stata incentrata sull’apprendimento delle funzionalità di SDL Trados, uno dei CAT Tools più utilizzati e più completi attualmente sul mercato. Per l’organizzazione dell’evento è stata stipulata una partnership con SDL, che ha concesso di svolgere presso la sede dell’università di Palermo l’esame per il conseguimento della certificazione di primo livello<sup>32</sup> di competenza nell’utilizzo della suite di prodotti Trados. I laboratori pratici sono stati quindi fortemente segnati da questa finalità, pur tentando al tempo stesso di presentare il software in maniera funzionale e concreta, tramite una modalità *hands-on* che ha dato a tutti gli studenti la possibilità di fare un’esperienza diretta del *tool* sia in maniera guidata che liberamente, dando spazio all’intuito e all’auto-formazione. Tuttavia, soprattutto nel corso della prima edizione della Summer School, l’obiettivo della certificazione è stato primario, e ha determinato l’approccio didattico in maniera netta. Ne è una prova il fatto che degli studenti che hanno preso parte alla prima edizione, più dell’80% ha conseguito la certificazione di primo livello alla fine del corso, e di questi circa il 60% con un punteggio molto elevato<sup>33</sup>. Un dato altrettanto importante e in controtendenza riguarda il numero esiguo (circa il 15%) di studenti che hanno deciso di investire economicamente sull’acquisto – per altro nettamente scontato – del pacchetto software. Se da un lato quindi si è dimostrato grande interesse nell’ottenere una certificazione, dall’altro tale interesse non è corrisposto all’intenzione di continuare ad utilizzare le conoscenze acquisite anche dopo il corso.

Durante la seconda edizione (2009), che ha ricalcato in buona parte l’organizzazione della prima e a partire dalla quale chi scrive è stata impegnata in qualità

---

<sup>31</sup> Nello specifico, la Metropolitan Leeds University (Regno Unito), la Politehnica di Timisoara (Romania) e la Salzburg Universität (Austria)

<sup>32</sup> La certificazione SDL Trados si articola in tre livelli di competenza ([www.sdl.com/services/education-certification/certificationproduct/sdl-trados-certification](http://www.sdl.com/services/education-certification/certificationproduct/sdl-trados-certification))

<sup>33</sup> L’esame consiste in un quiz a risposta multipla composto da 4 sezioni da 10 domande ciascuna, a cui dare risposta entro un tempo massimo di 2 minuti. Circa il 65% degli studenti ha totalizzato un punteggio pari o superiore a 36.

di tutor nei laboratori pratici di Trados, la maggiore novità è stata rappresentata dall'introduzione della preparazione, per quegli studenti che avevano partecipato l'anno precedente, al secondo livello di certificazione. Si sono quindi formate due classi, rispettivamente per la preparazione all'ottenimento dei due livelli. Avendo personalmente ottenuto la certificazione di secondo livello precedentemente alla Summer School, la scrivente ha seguito il gruppo degli studenti che per il secondo anno si trovavano ad approfondire le loro conoscenze sui CAT Tools. Il gruppo, di numero ridotto rispetto a quello dei principianti, era essenzialmente formato da quegli studenti che si erano dimostrati maggiormente interessati l'anno precedente, e molti dei quali avevano acquistato il software, continuando ad esercitarsi nel corso dell'anno. Questa combinazione di fattori – l'interesse personale dei partecipanti, la loro pregressa esperienza, il numero ridotto – ha permesso di organizzare il corso in maniera diversa rispetto all'anno precedente. Sebbene l'importanza della certificazione aleggiasse sempre su tutte le lezioni, l'approccio degli studenti, forti dell'anno precedente, è stato maggiormente incentrato sull'utilità del *tool* come strumento di lavoro e come competenza vincente per proporsi sul mercato. Si sono organizzate delle esercitazioni – simulazioni, sia per sperimentare la divisione del lavoro tra traduttori e revisori, con un rapporto uno a uno che favorisse il confronto e l'apprendimento, sia per provare una modalità di lavoro condiviso, con l'utilizzo di un'unica TM in linea e la gestione dei problemi di traduzione che coinvolgono l'uso attivo del CAT Tool, come l'aggiornamento e la modifica di voci salvate nella TM, la condivisione delle risorse e la divisione equa del lavoro sulla base dei calcoli di analisi del testo da tradurre. La competenza acquisita è stata così più personale e più concreta rispetto all'anno precedente. Anche la preparazione alla certificazione è stata vissuta più come metodo per approfondire l'uso del software piuttosto che in maniera fine a sé stessa. Il 100% degli studenti del corso avanzato ha conseguito la certificazione, pur notandosi un abbassamento del punteggio medio, che non ha mai raggiunto il massimo (40/40)<sup>34</sup>.

La terza edizione della Summer School si è inizialmente rivelata la più problematica a causa del subentro di una nuova versione di Trados (SDL Trados Studio 2010) con un'interfaccia e comandi nuovi rispetto alla versione precedente (SDL Trados

---

<sup>34</sup> Ciò è stato principalmente dovuto alla difficoltà oggettiva del test, maggiore rispetto a quello di primo livello.

2007), e di conseguenza una nuova certificazione. Ciò ha destabilizzato gli studenti che l'anno precedente avevano ottenuto una certificazione di primo livello e avrebbero voluto continuare col secondo livello. In particolare, essi hanno temuto che la certificazione ottenuta circa la versione precedente non fosse più valida (problema che in realtà non sussisteva data la permanenza sul mercato di entrambe le versioni del software e i tempi relativamente lunghi di agenzie e freelance di adattamento alle nuove versioni dei programmi). Il dover ricominciare da capo ha prodotto in alcuni un'iniziale frustrazione, principalmente dovuta alla consapevolezza acquisita circa il rapido processo di aggiornamento delle tecnologie e la conseguente inadeguatezza delle competenze pregresse. In realtà, non appena gli studenti alla seconda esperienza con Trados hanno sperimentato le funzioni della nuova versione, si sono potuti rendere conto che quanto già noto non era andato perduto, ma era servito come un prezioso background che ha reso l'apprendimento della nuova versione più rapido e agevole. Inoltre, potendo mettere a confronto le due versioni del software, gli studenti hanno potuto constatare che la seconda fosse per diversi aspetti di più facile approccio e di più intuitivo utilizzo. Grazie a tale confronto e alla consapevolezza acquisita di poter sfruttare la conoscenza di un software per cominciare ad utilizzarne altre versioni o tipologie, gli studenti si sono aperti alla possibilità di valutare lo strumento migliore da adottare in base alle loro preferenze o esigenze, o comunque hanno sviluppato un atteggiamento più aperto e fiducioso nei confronti dei CAT Tools.

#### **4.5.3. I laboratori di CAT Tools**

Oltre alle 3 edizioni della Summer School, negli anni successivi sono stati svolti dei laboratori di sussidi per la traduzione come parte pratica all'interno dell'insegnamento di Teoria e Tecnica della traduzione in corsi di laurea triennale. Chi scrive ha personalmente guidato per due anni consecutivi un totale di quattro laboratori, con una partecipazione media di circa 20 studenti per classe. Questi corsi, della durata di circa 25 ore, non essendo indirizzati all'ottenimento di alcuna certificazione hanno consentito di svolgere un lavoro più libero, che ha beneficiato comunque dell'esperienza maturata nei tre anni precedenti. È stato durante tali laboratori che si è implementata una modalità sperimentale di approccio ai CAT Tools, in cui sono state testate sugli studenti – tutti alla loro prima esperienza con queste tipologie di software – alcune delle proposte di applicazione didattica presentate nel corso di questo capitolo. Adottando i criteri

dell'approccio costruttivista, il corso ha inizialmente fornito agli studenti le informazioni necessarie a contestualizzare la tipologia di strumenti (distinti dai programmi di traduzione automatica), le loro finalità e potenzialità, per lasciare poi ampio spazio all'utilizzo pratico guidato ma al contempo con ampi spazi di autonomia per gli studenti al fine di incentivare un approccio personale e una sperimentazione libera del CAT Tool. Utilizzando dei testi con livelli di complessità crescente, sono state acquisite le funzioni fondamentali e alcune funzioni avanzate utili ad un uso competente e completo del software. Una volta raggiunto un certo grado di familiarità con il CAT Tool, si è data la possibilità ad ogni studente di scegliere un ambito specifico (ad es. sport, moda, ingegneria, elettronica, turismo, medicina) e raccogliere materiali bilingui da siti web attendibili per la costruzione delle prime TM. Ciò ha permesso loro di sviluppare la capacità di valutazione dell'affidabilità delle fonti, oltre che di costruire tramite allineamento e creazione di glossari il loro primo bagaglio di risorse per il lavoro. La ricerca dei materiali su internet ha avuto la sua parte di importanza, essendo molto utile per un traduttore sapere dove ricercare le informazioni necessarie e conoscere le risorse a cui può avere accesso gratuitamente sul web (cf. 2.5). Anche all'interno di questi laboratori si sono proposte delle esercitazioni composite in cui si è simulata la divisione di un progetto tra più traduttori e altrettanti revisori, e si sono invitati gli studenti a riflettere sia sull'apprendimento linguistico (che non ha comunque costituito uno degli obiettivi principali del corso) che sulla preparazione pratica in vista di una possibile esperienza lavorativa.

Nel corso del laboratorio, si sono potute appurare alcune reazioni significative degli studenti, carpite osservando il loro comportamento nel corso delle esercitazioni, o stimolate tramite espressa richiesta di giudizi, commenti e percezioni. In primo luogo, è stata riconosciuta con generale consenso l'utilità dello strumento, in particolare del contributo positivo di supporto alla traduzione e di organizzazione del lavoro. Tuttavia, per un esiguo numero di studenti (ammontanti al 3% dei partecipanti) che hanno manifestato di avere un rapporto difficile con le tecnologie in generale, l'utilizzo dei CAT Tool è stato visto come una fonte di stress e un possibile ostacolo nello svolgimento di traduzioni in un contesto reale. L'ansia di non utilizzare correttamente i comandi e la sensazione di complicazione piuttosto che di vantaggio ha fatto sì che essi sviluppassero un'attitudine negativa verso tali strumenti, che ha compromesso l'intero percorso di

apprendimento<sup>35</sup>. Al contrario, si è potuto osservare in certi casi un forse eccessivo entusiasmo, e la tendenza a confidare in maniera assoluta sulle capacità dei sistemi digitali, come a demandare interamente ad essi il compito di ricordare quanto già tradotto. Ne sono prova il fatto che in diversi casi gli studenti non erano in grado di evocare a mente la traduzione di un termine incontrato diverse volte nelle traduzioni, nonché un scarso livello di pensiero critico nell'adozione dei suggerimenti desunti dalla TM, che inficiavano la qualità finale della traduzione. A tal riguardo si è cercato di correggere il tiro durante il corso, invitando gli studenti a riflettere sulla loro responsabilità in quanto traduttori e sui limiti da riconoscere ai software CAT. Allo stesso tempo è da mettere in luce l'approccio di alcuni studenti che si sono detti interessati ad avviare una carriera nel settore della traduzione. Oltre a dimostrarsi più attenti e più motivati circa l'apprendimento del *tool*, essi hanno anche saputo cogliere gli inviti a vedere la formazione pratica in maniera integrata alla formazione di tipo linguistico, approfittando delle esercitazioni anche per apprendere nuove espressioni e nuova terminologia.

A conclusione del laboratorio, la prova finale è consistita in una breve traduzione svolta con l'ausilio del CAT Tool, e da un breve colloquio finalizzato a cogliere il grado di comprensione e consapevolezza delle conoscenze acquisite, il tipo di processo seguito durante lo svolgimento di una traduzione con un CAT Tool e il grado di soddisfazione e percezione dell'utilità dello strumento. Dalla verbalizzazione delle esperienze individuali si è potuta ricavare un'opinione comune positiva, con gradi diversi di profondità e di visione in prospettiva futura. Diversi studenti hanno riconosciuto tra gli apporti positivi del CAT Tools la possibilità di seguire il lavoro in maniera più ordinata e metodica, mentre gli studenti più coinvolti hanno parlato di un'utilità a lungo termine per la raccolta e l'organizzazione funzionale di materiali utili per il lavoro. Tra le difficoltà maggiori riscontrate, la necessità di adattare il proprio *modus operandi* al metodo "imposto" dall'uso del software, e la necessità di gestire la quantità e le molteplici tipologie di file e cartelle create durante il processo.

Il corso sicuramente non è stato sufficiente a rendere gli studenti pienamente competenti né ad utilizzare in maniera ottimale un CAT Tool. È stato tuttavia sufficiente ad accendere una curiosità e a fornire le conoscenze basilari per poter approfondire il campo. Inoltre,

---

<sup>35</sup> Questa tendenza trova conferma in Biau Gil (2006: 94), in cui viene evidenziato come il livello di partenza di conoscenze informatiche influenzi il grado di autonomia degli studenti e la loro capacità di utilizzare le risorse in maniera più competente.

la sperimentazione di un metodo di più ampio respiro ha sofferto di una certa ingenuità data dall'ancora debole base teorica da parte della scrivente in qualità di docente, base che è stata rafforzata in un secondo momento, durante gli anni di studio dottorale e per dinamiche interne all'università corrispondenti alla momentanea cessazione di attività laboratoriali nel settore delle tecnologie per la traduzione. L'intenzione è di poter riprendere nel prossimo futuro l'approccio sperimentale avviato, per ottenere dati più significativi relativamente all'ipotesi di formazione professionale e linguistica attraverso l'adozione dei CAT Tools.

## PARTE SECONDA

### LA TRADUZIONE AUTOMATICA

## Capitolo 5: La storia e i modelli

### 5.1 *Una introduzione*

#### 5.1.1. **Un argomento controverso**

Dopo aver interamente dedicato la Prima parte di questo lavoro ai CAT Tools, sufficiente spazio verrà ora lasciato alla discussione sulla traduzione automatica (TA) o in inglese *machine translation* (MT), un argomento per molti aspetti controverso sul quale si è sviluppato un lungo e vivace dibattito che vede al suo interno posizioni molto diverse e persino contrastanti. In effetti, fin dal suo esordio, tale fenomeno è stato accompagnato da questa dicotomia, espressa da dichiarazioni ufficiali, studi e propagande dai toni ora entusiastici, ora scettici. Furono anche formulate definizioni diverse di *machine translation*, alcune delle quali persino contraddittorie tra di loro. Originariamente, il termine si riferiva esclusivamente ai sistemi automatici che non contemplavano alcun intervento umano (Sager, 1994; Weaver, 1949). La *European Association of Machine Translation* (EAMT) definisce questa attività come “*the application of computers to the task of translating texts from one natural language to another*” (cit. in Quah, 2006: 9), definizione alla quale fa eco quella dell’*International Association of Machine Translation* (IAMT), che sottolinea inoltre la modalità di elaborazione frase per frase dei testi (*ibid.*). Infine, in Arnold *et al.* (1994:1), così come in altri studi (Melby, 1981; Arnold, 2000; Hutchins 2001) si fa invece riferimento ad un possibile intervento umano affermando che la traduzione automatica è “il tentativo di automatizzare del tutto o in parte il processo traduttivo da una lingua ad un’altra”. Forcada (2010) chiama il prodotto della traduzione automatica “*raw translation*”, a sottolineare la necessità di sottoporre tale prodotto alla revisione da parte di un professionista. La *machine translation* è in effetti un’attività e un ambito di studio articolato, nei confronti del quale gli studiosi si avvicinano in maniera diversa a seconda delle loro personali convinzioni e del loro background formativo. In particolare, tre tipologie di esperti contribuiscono, ognuna nella maniera che le è propria,

allo studio di questo fenomeno: gli sviluppatori dei software, i ricercatori nell'ambito delle intelligenze artificiali (ingegneri, matematici ma anche linguisti), e infine i traduttori o studiosi di traduzione (Pym, 2011b: 2). Gli esperti e ricercatori appartenenti a queste categorie studiano il fenomeno da prospettive e con finalità diverse, lo valutano secondo criteri diversi e ne parlano in maniera circoscritta al proprio ambito di interesse. È normale, quindi, che vi siano dichiarazioni e punti di vista tanto differenti da giungere ad essere contrastanti. Tuttavia, al di là delle convinzioni soggettive o dei diversi approcci, è innegabile che il “fenomeno traduzione automatica” si sia talmente diffuso e affermato da non poter essere ignorato, o da non poter trattare in maniera marginale il grande impatto prodotto sui campi di ricerca afferenti ad esso, sull'industria della traduzione e sul modo in cui molte persone si relazionano tramite di esso alle lingue straniere.

Le motivazioni dietro all'affermazione della traduzione automatica sono molteplici. Alcune sono di natura essenzialmente pratica, come il bisogno di poter accedere rapidamente al contenuto di testi scritti in lingue che non si conoscono o l'esigenza di ottenere traduzioni a basso costo o persino gratuitamente, seppur a discapito della qualità del risultato. Vi sono tuttavia anche delle motivazioni di natura più speculativa, come il desiderio degli studiosi di esplorare e riprodurre il processo traduttivo umano e le dinamiche comunicative proprie dei linguaggi naturali, e di sperimentare fino a che punto possano spingersi le intelligenze artificiali nello svolgimento di operazioni non numeriche. La possibilità di munire le macchine delle competenze linguistiche necessarie ad interloquire con l'utente in una o più lingue è stata una chimera che ha accompagnato lo sviluppo dei computer fin dall'esordio dei primi rudimentali modelli, insieme allo studio e alla sperimentazione di modalità e processi che consentissero la comprensione, produzione e traduzione di testi da parte del computer senza alcun intervento umano. Queste ambiziose teorie, sinora mai realizzate fino in fondo, caricarono gli esordi della traduzione automatica di un'aspettativa e un ottimismo decisamente superiori alle sue reali potenzialità, inducendo molti a credere alla reale possibilità di avere in pochi anni dei computer capaci di elaborare e produrre linguaggio correttamente. I primi risultati reali, e qualche voce fuori dal coro (Kay, 1980) rivelarono il bisogno di aggiustare il tiro delle sperimentazioni e rivalutare i reali utilizzi possibili di tale strumento.

### 5.1.2. Le voci contrarie

Oltre che da tale grande e persino sovra misurato entusiasmo, il fenomeno della traduzione automatica è stato accompagnato fin da subito da una forte opposizione, anch'essa sostenuta da motivazioni sia pratiche che di principio. A dar voce a tale opposizione erano e sono tutt'ora principalmente traduttori e studiosi "canonici", insoddisfatti dai risultati scadenti forniti dai computer e preoccupati per il "danno di immagine" arrecato alla professione di traduttore dall'ipotizzata possibilità di far svolgere tale attività ad una macchina. I sostenitori della traduzione come arte (Holmström 1951 cit. in Hutchins 2001) non possono accettare di chiamare "traduzione" quella compiuta in maniera automatica dal computer, soprattutto perché la tipologia testuale alla quale essi fanno più spesso riferimento è anche la meno adatta ad essere sottoposta a traduzione automatica, ovvero la letteratura. Dal loro canto, invece, i traduttori professionisti vi si oppongono da un lato per il timore che la TA possa in alcuni casi sostituirsi ad essi privandoli del proprio lavoro, dall'altro per il modo in cui l'adozione tattica della traduzione automatica da parte delle agenzie di traduzione abbia trasformato in molti casi il loro lavoro in una snervante e frustrante revisione di testi tradotti automaticamente, mal formulati e sconnessi. Tale revisione tra l'altro viene solitamente pagata meno della traduzione vera e propria. L'astio dei traduttori verso la traduzione automatica è antico, risale già ai primi anni della sua comparsa, ed è per molti versi legata al fatto che, nello sviluppo di tale strumento da parte dei diversi centri di ricerca, essi non fossero mai stati chiamati in causa per offrire la loro esperienza e professionalità, e modellare sulla base dei loro bisogni e delle loro aspettative una risorsa che era destinata anche a loro (Hutchins, 2003; Pym, 2011a). L'essere esclusi dalla ricerca portò i traduttori ad assumere un'attitudine negativa e oppositiva nei confronti della traduzione automatica, nel timore che essa avrebbe potuto danneggiarli. In effetti, in diversi casi la traduzione automatica ha una ricaduta concreta sul lavoro dei traduttori, anche se, ad oggi, nessuno metterebbe in dubbio il fatto che la traduzione automatica e quella umana siano due processi diversi e non comparabili dal punto di vista della qualità e dell'affidabilità. Gli interessi economici, tuttavia, hanno spesso la meglio sugli altri aspetti, e capita che la prima venga adoperata laddove invece sarebbe necessario avvalersi della seconda. Sempre più si è iniziato a comprendere l'utilità della complementarità delle due tipologie di traduzione, e il beneficio che deriva dal saper adottare ora l'una ora l'altra in base al caso specifico,

alla tipologia testuale da trattare ed al tipo di risultato finale che si desidera ottenere. La validità dei sistemi di TA deve infatti essere sempre valutata all'interno di un contesto di utilizzo e nella giusta prospettiva, con la consapevolezza dei suoi vantaggi e demeriti.

### **5.1.3. Prospettive diverse**

La traduzione automatica è un argomento importante anche da punti di vista diversi da quello linguistico. Essa ha infatti implicazioni sociali, commerciali, economiche, scientifiche e filosofiche (Arnold *et al.*, 1994: 5). Le implicazioni sociali sorgono laddove ci si soffermi a pensare al fatto che la TA si è evoluta in risposta ad un bisogno crescente di traduzioni in tutto il mondo, frutto di cambiamenti sociali ed economici epocali. È un bisogno di traduzioni che i traduttori umani non riescono a gestire senza un aiuto tecnologico, e che non sempre tra l'altro necessita del loro intervento, essendo sufficiente l'input seppur imperfetto della macchina, come si vedrà più avanti (cf. par. 7.2). Inoltre, l'affermazione della traduzione automatica ha cambiato il modo degli utenti comuni (non traduttori) di approcciarsi alle lingue straniere, fornendo a tutti una chiave di accesso facile e veloce a contenuti altrimenti inavvicinabili, pur con i limiti già citati relativi alla qualità della traduzione ottenuta. Rientra nelle ricadute sociali di questo fenomeno anche l'impatto sulla professione dei traduttori, chiamati in un modo o nell'altro a fare i conti con il loro rivale/collaboratore digitale, e sul costo della traduzione come già accennato, specialmente per istituzioni come ad esempio l'Unione Europea e società come le multinazionali che hanno continuo bisogno di traduzioni. L'impatto scientifico è relativo all'avanzamento dei settori tecnologici e linguistici che hanno camminato di pari passo allo sviluppo della TA, come le scienze informatiche e le intelligenze artificiali, o la linguistica computazionale e il *Natural Language Processing*, mentre quello filosofico riguarda essenzialmente il processo di formulazione del linguaggio e il tentativo di automatizzare un'attività che richiede l'intera gamma delle conoscenze umane, quale è la comunicazione. A questi aspetti si dedicherà spazio all'interno del settimo capitolo.

Nel corso di questa Seconda Parte, oltre a seguire l'evoluzione storica della traduzione automatica e a descrivere i diversi sistemi che si sono avvicinati dalle sue origini ad oggi, si cercherà di scendere più in profondità nell'analisi delle teorie sulle quali tali sistemi si basano, e di cogliere le diverse concezioni del linguaggio e in particolare del processo traduttivo umano come attività cognitiva, soggiacenti allo

sviluppo dei vari modelli di TA. A partire da ciò si tenterà di sviluppare una riflessione critica sull'argomento, non finalizzata a sostenere l'una o l'altra posizione nel dibattito sulla fattibilità di una traduzione interamente automatica. Piuttosto, si cercherà di comprendere fino a che punto tale strumento riesca a riprodurre i meccanismi mentali legati al linguaggio e alla traduzione, e in un secondo momento quale tipo di aiuto può effettivamente venire dai sistemi di TA per i traduttori e per gli utilizzatori non specialisti di tali strumenti. A tal fine si farà riferimento a numerosi studi, alcuni dei quali hanno rappresentato delle svolte storiche nell'evoluzione del fenomeno, e altri che presenteranno invece la situazione attuale e gli ultimi sviluppi. In particolare, nel presente capitolo verranno presentati o citati specifici sistemi sviluppati in ambito universitario e di ricerca – quindi sperimentali e dimostrativi – o implementati da società di servizi linguistici o di *information technology* quali sistemi a pagamento, spesso rivolti a clienti specifici o settori ben determinati del mercato. Non si parlerà in questo momento degli odierni sistemi di traduzione automatica disponibili in maniera gratuita sul web, come i celebri Google Translate o Bing Translator; ad essi verrà dedicato spazio al capitolo sesto.

## 5.2 *Sistemi di prima generazione*

### 5.2.1. **I primi studi**

La traduzione automatica è un fenomeno nato da studi di natura interdisciplinare, che traevano spunto dalla linguistica tradizionale e dalla teoria della traduzione, ma anche dalla linguistica informatica, la statistica e le intelligenze artificiali. L'ipotesi che diede il via al suo primo sviluppo tra gli anni '40 e '50 del secolo scorso sosteneva che le lingue naturali potessero essere esaurientemente descritte, controllate e codificate secondo la stessa logica che regola le operazioni matematiche, quindi riprodotte con facilità da un dispositivo digitale (Wills, 1999, cit. in Quah 2006). La traduzione diventava così secondo tale prospettiva un mero processo di decodifica di un linguaggio e di successiva ricodifica del contenuto in un secondo linguaggio o "codice"<sup>36</sup> (Weaver, 1955). Tale trasposizione, vista come un processo più matematico che linguistico, quindi fondato su

---

<sup>36</sup> Nel 1955 scriveva Weaver a Booth, "It is very tempting to say that a book in Chinese is simply a book written in English which was coded into the 'Chinese code'. If we have useful methods for solving almost any cryptographic problem, may it not be that with proper interpretation we already have useful methods for translation?" (cit. in Stein, 2013: 6)

regole esatte, fisse e univoche, poteva essere compiuta da una macchina sufficientemente equipaggiata con le informazioni necessarie allo svolgimento di tale compito.

Un breve accenno sugli esordi della traduzione automatica è già stato fatto all'inizio del primo capitolo (cf. 1.2), essendo questo fenomeno legato strettamente allo sviluppo dei CAT Tools e persino antecedente ad essi. Sebbene già in tempi più antichi fossero state sviluppate delle teorie filosofiche sul linguaggio che possono essere considerate le antesignane della ricerca sull'automazione linguistica<sup>37</sup> (Stein 2013: 5), la sua prima manifestazione avvenne attorno alla metà del secolo scorso, principalmente in ambito militare e di intelligence, allo scopo di poter rapidamente e agevolmente decifrare i messaggi della fazione avversaria durante la guerra fredda tra Stati Uniti e Unione Sovietica, non a caso i primi due stati ad avere investito grandi cifre sulla sperimentazione relativa alla traduzione automatica. Può apparire per certi versi paradossale, se si pensa allo sviluppo ancora immaturo dei computer di allora e alle loro limitate capacità di elaborazione dei dati, che si sia pensato subito non tanto di utilizzarli come supporto al processo traduttivo umano, ma piuttosto interamente come suoi sostituto che dovessero operare in autonomia. In realtà, l'esigenza di ottenere traduzioni in tempi molto ristretti (se non addirittura istantaneamente) e con costi decisamente contenuti, rese l'idea della traduzione automatica molto appetibile. Inoltre, il fascino delle intelligenze artificiali e della presunta capacità del cervello elettronico di saper apprendere ed utilizzare il linguaggio umano accese l'entusiasmo degli studiosi e diede vita a un fiorente studio con numerose sperimentazioni e approcci diversi fino ai giorni nostri. L'immagine seguente (Quah 2006: 58) mostra il succedersi dei principali approcci adottati dagli anni '50 fino ad oggi nello sviluppo e nello studio dei sistemi di traduzione automatica:

---

<sup>37</sup> Stein (2013) fa riferimento in particolare alla teoria delle monadi di Leibniz (1696), nella quale i "termini primi" erano le più piccole unità di pensiero che potevano rappresentare mentalmente tutta la conoscenza ed il pensiero umano, e alla proposta avanzata dal chimico tedesco Becher (1661) di una tecnica crittografica molto simile a quella proposta da Weaver secoli dopo, grazie alla quale si potevano comprendere "diverse lingue se non persino tutte".

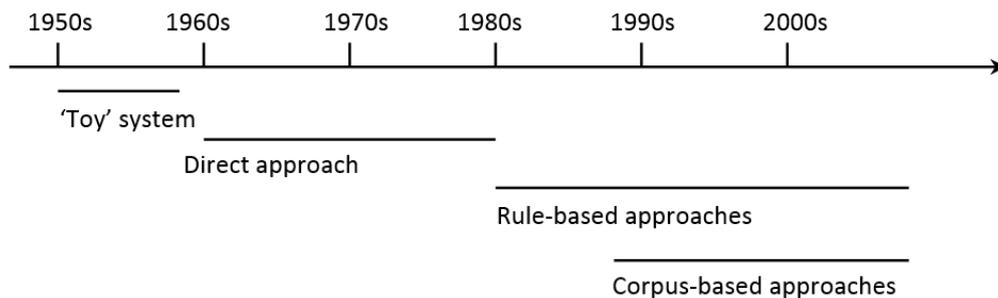


Fig. 25: Cronologia degli approcci nella traduzione automatica (Quah, 2006: 58)

### 5.2.2. I sistemi *direct-approach*

Il primo a parlare di traduzione automatica fu, nel 1949, il già citato Warren Weaver, allora direttore della Rockefeller Foundation, che con il suo celebre memorandum segnò l'inizio della ricerca in questo ambito, ponendo l'accento su alcuni aspetti peculiari come la logica del linguaggio, i suoi elementi invariabili e infine le tecniche di crittografia, che egli proponeva di utilizzare per la meccanizzazione del processo traduttivo. Tale approccio fu già allora criticato da molti per la sua idea semplicistica e persino ingenua di linguaggio, cosa tra l'altro riconosciuta dallo stesso Weaver nell'introduzione al memorandum. Malgrado tali critiche, esso accese comunque l'interesse dei ricercatori, tanto da suscitare l'avvio di sperimentazioni sia in Inghilterra col progetto Colossus, che negli Stati Uniti con il progetto ENIAC sviluppato dall'esercito. Sfortunatamente ma anche prevedibilmente, le tecniche crittografiche non si rivelarono utili al fine di decifrare e tradurre testi in linguaggio naturale, per via della scarsa prevedibilità e logicità matematica delle lingue. Tuttavia queste prime sperimentazioni produssero un serio interesse nei confronti dei processi di meccanizzazione della traduzione, spingendo l'agenzia di intelligence statunitense, la CIA, e quella sovietica, il KGB, a investire ingenti capitali su tale ricerca. Dagli studi condotti in questo periodo nacquero i così detti sistemi di "prima generazione" o *Direct – Approach MT Systems*, basati su una forte preponderanza della componente lessicale e da un tentativo di corrispondenze parola per parola (*word-based MT*), nonché sull'ambizioso perseguimento di una traduzione FAHQMT (*fully automatic high quality machine translation*). Il primo sistema reso pubblico fu prodotto in collaborazione da IBM e dalla Georgetown University (Hutchins 2005: 2); si trattava di un traduttore automatico russo – inglese dalle caratteristiche estremamente elementari tanto da essere

considerato poco più di un “sistema giocattolo” (Hutchins 1995: 440). Esso infatti riconosceva solamente 250 parole, 6 regole grammaticali e 49 frasi, eppure bastò a far stanziare al governo statunitense fondi per circa 20 milioni di dollari (Somers 2003: 4) per lo sviluppo di sistemi più articolati e di maggiori dimensioni, che fornissero traduzioni interamente automatiche e di qualità. Frutto di queste ricerche furono dei sistemi di gestione terminologica quali ABC di ALPS, Termex, prodotto da LinguaTech, e INK TextTools (Balkan, 1992), i quali consentivano agli utenti di consultare dizionari e glossari elettronici precompilati, e di crearne di personalizzati contenenti oltre alla traduzione dei termini anche informazioni e riferimenti utili. I risultati, visualizzati all’interno della schermata del programma di videoscrittura, potevano essere estratti e inseriti direttamente nel testo, o modificati in corso d’opera. INK inoltre era anche in grado di fornire dati sulla frequenza d’uso dei termini e alcune informazioni sul contesto. Queste caratteristiche si ritroveranno ampliate, migliorate e arricchite anche all’interno dei software più moderni. Occorre tuttavia specificare che tali strumenti, essendo *language-dependent*, erano in grado di elaborare un numero molto limitato di lingue, spesso solo due, a causa della scarsa capacità interna dei computer di quel periodo, quindi dell’impossibilità di salvare nella loro memoria numerosi vocabolari e grammatiche per un numero superiore di lingue. Un altro ostacolo era rappresentato proprio dalla programmazione dei sistemi di TA, che dovevano essere predisposti in fase preliminare all’utilizzo di una data lingua, con la definizione di un set di regole grammaticali e sintattiche per la scomposizione del testo in entrata e la riformulazione dei segmenti di testo identificati nella lingua di arrivo e secondo il suo set di regole.

### **5.2.3. I sistemi *phrase-based* e il linguaggi controllati**

I nuovi sistemi di traduzione automatica, sempre afferenti alla prima generazione ma più complessi rispetto all’esperimento della Georgetown, furono invece ALPS, LOGOS e Systran (Balkan, 1992). Tali sistemi utilizzavano ancora il metodo della “traduzione diretta”, basato sull’idea che la versione nella lingua target potesse generarsi da una corrispondenza parola per parola con il testo originario, ma univano a questo anche un tentativo di analisi sintattica. Per questo ci si riferisce a loro come modelli *phrase-based* o *syntax-based* (Stein, 2013: 10), non tanto perché tenessero in conto la frase come unità minima di analisi e traduzione, ma piuttosto per il loro tentativo di tradurre parole insieme, sulla base dell’analisi compiuta. I dati sintattici ottenuti potevano risolvere

problemi di riordino delle parole nella frase tradotta, o ridurre le possibili ambiguità insorgenti dall'identificazione erronea delle parti del discorso (nomi, verbi, aggettivi ecc.).

ALPS, prodotto da Alpnet, prima società di servizi linguisti degli Stati Uniti, racchiudeva in sé sia le potenzialità dei sistemi di gestione terminologica che la capacità di produrre una bozza di traduzione frase per frase, che necessitava tuttavia dell'intervento del traduttore umano ad ogni passaggio. LOGOS era un programma più sofisticato, basato sul confronto del testo con dizionari e database terminologici di riferimento, compilati con l'aggiunta di informazioni linguistiche e semantiche al fine di produrre una traduzione automatica che richiedesse l'intervento umano solo per accettare o rifiutare le traduzioni suggerite. La velocità di produzione era di 1000 parole all'ora. Systran (System Translation), prodotto dal ricercatore americano Peter Toma e utilizzato dall'aeronautica statunitense, includeva le stesse funzioni dei precedenti sistemi, ma con una maggiore velocità di traduzione (150.000 parole all'ora) e la divisione delle componenti linguistiche e computazionali in due moduli separati al fine di non creare interferenze tra di esse. Fu inoltre introdotta una nuova funzione, rappresentata dal tentativo di traduzione delle parole sconosciute sulla base dell'osservazione delle regole morfologiche.

Questi tre sistemi furono principalmente applicati ad ambiti linguistici molto ristretti, o più precisamente a linguaggi controllati, caratterizzati da un vocabolario, una grammatica e uno stile artificialmente semplificati al fine di rendere i contenuti chiari, diretti e privi di possibili ambiguità, quindi anche più facili da tradurre automaticamente ottenendo risultati accettabili (Mitamura e Nyberg, 1995). I linguaggi controllati si basano su vincoli applicati al vocabolario di partenza, per ridurre le ambiguità e la complessità lessicale, e vincoli sulla struttura delle frasi, affinché la suddivisione da parte del software risulti più facile sia da identificare che da riprodurre. Alcune delle regole fondamentali su cui tali linguaggi sono costruiti sono l'utilizzo di frasi brevi che esprimano ognuna un unico concetto, l'adozione esclusiva della terminologia autorizzata e la sostituzione di espressioni ambigue, idiomatiche o polisemiche con altre univoche e il più possibile precise, la non omissione di alcuna parte del discorso, come nomi, articoli ecc. L'applicazione di queste regole, oltre a migliorare la qualità della traduzione automatica, rendeva i testi di partenza più facili da leggere e comprendere. Per questo motivo i linguaggi controllati si diffusero in molti settori per la redazione della manualistica o di

documenti tanto divulgativi quanto tecnici. È il caso dell'inglese semplificato per il settore automobilistico (AECMA), i sistemi Caterpillar sviluppati dall'università di Carnegie-Mellon ed utilizzati dalla Xerox Corporation ecc. (Hutchins 2005: 9). I sistemi di TA operanti con tali linguaggi controllati erano quindi fundamentalmente “*custom-built*” (Hutchins 2003: 12), ovvero sviluppati appositamente per uno specifico cliente o un settore ben delimitato, o persino un singolo prodotto. Un ambito linguistico così ristretto, e la fruizione di risorse di riferimento altamente settoriali e specifiche, erano in grado di fornire traduzioni utili e adeguate agli scopi perseguiti. L'esempio più celebre è rappresentato dal programma canadese TAUM Météo, sviluppato nel 1976 per tradurre bollettini meteorologici dall'inglese al francese e ancora oggi in funzione. Il suo vocabolario limitato e l'elevata ripetitività dei costrutti consentono risultati accettabili di traduzione (80% di traduzioni corrette), ed esso rappresenta uno dei sistemi automatici di maggiore successo mai sviluppati. Molti altri sistemi “*custom-built*” sono stati prodotti attorno agli anni '70 e sono ancora in funzione, come il sistema TITUS per la traduzione di documenti afferenti all'industria tessile da e verso l'inglese, il francese, il tedesco e lo spagnolo, o il più recente sistema Kielikone originariamente sviluppato per la Nokia e poi diffuso tra altre società finlandesi (Hutchins, 2003: 13). Diverse aziende e organizzazioni svilupparono dei sistemi di TA *in-house*, specifici per i loro bisogni e i tipi di prodotti commercializzati. In molti altri casi, piuttosto che sviluppare dei sistemi specifici o formulare degli appositi linguaggi controllati, divenne prassi comune introdurre una fase di pre-editing dei testi più complessi, atta a semplificarne la sintassi e la terminologia e ad estrarre solo i concetti fondamentali del testo, espressi in frasi semplici e lineari, proprio al fine di sottoporre i testi al programma di traduzione automatica. Come sottolineato da Andreyewski (1981), la traduzione automatica auspicata e perseguita dai ricercatori era essenzialmente un'attività robotica, delegata interamente alla macchina. Come tale, essa poteva prevedere ed elaborare solo dati chiari, lineari, il più possibile univoci e di numero finito, motivo per il quale essa poteva solo applicarsi a testi adatti alle capacità robotiche del sistema. Tale processo di semplificazione testuale o il ricorso ai linguaggi controllati, tuttavia, si rivelò costoso e poco conveniente per la qualità comunque non ottimale dei risultati ottenuti, nonché per il ricorso all'intervento umano che veniva ad eliminare il vantaggio ricercato della totale automazione del processo.

#### 5.2.4. Il report ALPAC

Nel 1966 nacque *l'Automatic Language Processing Advisory Committee* (ALPAC), ente preposto alla verifica della qualità dei sistemi di traduzione automatica sino ad allora prodotti. Il report da esso pubblicato in quell'anno concludeva che i sistemi erano lenti, poco accurati e costosi, ben lungi dagli obiettivi che i ricercatori si erano prefissati. Secondo i commissari dell'ALPAC non vi erano inoltre possibilità di un sostanziale miglioramento nei risultati ottenibili, che giustificasse un ulteriore avanzamento degli studi con altro stanziamento di fondi<sup>38</sup>. Piuttosto, il report suggeriva di concentrare le ricerche su strumenti di sussidio alla traduzione e sull'approfondimento della linguistica computazionale. Il sentire generale di studiosi ed enti governativi passò rapidamente da un acceso ottimismo ad un fatalistico abbandono di tali sistemi, e per qualche tempo la traduzione automatica fu accantonata, per lo meno negli Stati Uniti. Altri studi furono avviati in Canada, in Germania e nei paesi dell'est, dall'Unione Sovietica fino alla Cina e al Giappone, con la produzione di sistemi di traduzione automatica bilingui. In Europa, attorno al 1970, la ricerca sulla traduzione automatica produsse EUROTRA (European Translation), un progetto franco-tedesco che includeva tutte le lingue allora parlate all'interno della Comunità Europea. Più che per la qualità dei suoi risultati, tale sistema si rivelò importante per lo sviluppo di altre ricerche ed altre sperimentazioni durante il ventennio successivo ed oltre, essendo il più grande e uno dei più ambiziosi progetti fino ad allora prodotti.

Tra i principali sostenitori della fazione dei dubbiosi fu il filosofo e linguista Bar Hillel, che nel 1959 dichiarò l'impossibilità della FAHQMT, ovvero una traduzione interamente automatica e di alta qualità per qualsiasi tipologia di testi. Occorreva infatti, secondo lo studioso, rinunciare ad una delle caratteristiche contenute nell'acronimo (Niremburg 1996: 1; Melby, 1981: 25). La traduzione interamente automatica non avrebbe potuto avere un'alta qualità, una traduzione di ottima qualità non poteva essere prodotta da un processo interamente automatico. Laddove si volesse tentare di mantenere entrambe le caratteristiche, si sarebbe dovuta restringere la tipologia testuale da sottoporre alla macchina, e prediligere come già detto i linguaggi controllati. Bar-Hillel, tuttavia,

---

<sup>38</sup> "The committee believes that it is wise to press forward undaunted, in the name of science, but that the motive for doing so cannot sensibly be any foreseeable improvement in practical translation. Perhaps our attitude might be different if there were some pressing need for machine translation, but we find none" (Pierce *et al.* ALPAC report, 1966: 24)

non intendeva affermare l'impossibilità totale di attuare una traduzione automatica, quanto porre l'accento sulle reali potenzialità di tale strumento in modo da indirizzare la ricerca e le aspettative degli utenti verso la giusta direzione<sup>39</sup>. Malgrado tali raccomandazioni, nel paese in cui più si era sinora investito sulla traduzione automatica, gli Stati Uniti, per più di un decennio le sperimentazioni si interruppero o vennero incentrate su tematiche più promettenti come la linguistica computazionale o le intelligenze artificiali.

### 5.3 *Sistemi di seconda generazione*

#### **5.3.1. L'approccio funzionale**

A partire dagli anni '80 negli USA vi fu un ritorno di interesse nei confronti della *machine translation*, grazie a dei risvolti nella linguistica computazionale che consentirono di sviluppare degli approcci più sofisticati alla traduzione. Nacquero in questo periodo SPANAM ed ENGSPAN (Arnold *et al.*, 1994: 14), due sistemi di traduzione inglese-spagnolo e viceversa, ed altri programmi che rappresentarono la seconda generazione di software per la traduzione automatica. Essi furono caratterizzati dall'introduzione di una componente di intelligenza artificiale (AI) per una più profonda analisi delle componenti testuali, e dall'adozione di un duplice approccio linguistico, uno formale e uno funzionale. Il primo poneva maggiore enfasi sulla descrizione delle strutture morfologiche e sintattiche, il secondo si concentrava invece sull'utilizzo del linguaggio e sul modo in cui parole e frasi si combinano per produrre testi ben formati (Bennett cit. in Somers ed. 2003: 144). Il primo approccio, quello formale, risultò più facile da incorporare nei sistemi di traduzione automatica, e influenzò maggiormente gli studi e le sperimentazioni, soprattutto con la diffusione della grammatica generativa di Chomsky. L'obiettivo dell'approccio funzionale era infatti quello di definire un numero limitato di regole e principi generali sulla base dei quali avviene la formazione delle strutture grammaticali e quindi la produzione di frasi e testi in un dato linguaggio. Il metodo formale ricorre alla rappresentazione della frase secondo uno schema che ne scompone la struttura ed il significato, mostrando la sua struttura sintattica. Un esempio

---

<sup>39</sup> Di fatto, le considerazioni di Bar-Hillel diedero nuovo impulso alla ricerca sulle intelligenze artificiali, essendo stata riconosciuta da lui l'incapacità dei computer di comprendere il linguaggio per mancanza di conoscenze del mondo reale (Melby, 1981: 25).

di applicazione di questo nuovo modello fu Metal, prodotto dall'Air Force statunitense e dall'Università del Texas, e lanciato nel 1988 da Siemens, che lo pubblicizzò come il primo sistema "intelligente" di traduzione automatica (Balkan 1992: 415). Rispetto ai programmi precedenti, esso poteva contare sull'utilizzo delle più complesse tecniche linguistiche e computazionali appena descritte, producendo uno schema sintattico ad albero per analizzare le frasi e dividerle in livelli sulla base di regole grammaticali e sintattiche. Al momento di tradurre, il sistema iniziava con i livelli più bassi, come le frasi principali, passando ai livelli successivi in base alle regole desunte dal testo. In questo modo, i programmatori di Metal sostenevano di poter garantire un'interpretazione corretta del testo originario e una traduzione adeguata. Anche il sistema di aggiornamento dei dizionari era più avanzato, poiché il programma era capace di ottenere informazioni sul termine a partire da un'analisi morfologica della radice e dal confronto con termini simili già presenti in memoria. Inoltre, per la prima volta un sistema di traduzione offriva la possibilità di preservare il layout del testo originario nella traduzione.

Un altro sistema prodotto al di fuori degli USA è il software israeliano TOVNA, che al pari delle pretese di Metal vantava di essere il primo programma di traduzione in grado di "imparare" dagli esempi, raccogliendo ed elaborando dati non solo linguistici, ma anche matematici e logici grazie all'applicazione dell'intelligenza artificiale (Segal 1991). Il sistema registrava ogni passaggio del processo di traduzione, cosicché se al momento del post-editing l'utente rilevava un errore, poteva recuperare le informazioni registrate e risalire alle regole grammaticali e sintattiche applicate in quel preciso passaggio per capire cosa aveva causato l'errore, in modo da modificare tali regole o introdurre una variante o un'espressione specifica. Ciò implicava quindi il ricorso all'intervento umano a fianco dell'intelligenza artificiale. Anche TOVNA utilizzava uno schema sintattico ad albero, costruito allo scopo di risolvere le ambiguità del testo sulla base di rimandi ai livelli precedenti e successivi, atti a verificare il testo con le diverse variabili possedute dal programma.

### **5.3.2. Il nuovo approccio formalista**

Sebbene questo nuovo approccio abbia rappresentato un notevole passo avanti, l'applicazione della grammatica generativa si rivelò inadatta ad essere applicata alla traduzione automatica, poiché richiedeva una programmazione informatica ancora per i tempi eccessivamente complessa ed estesa. Venne così recuperato l'approccio formalista,

al quale però vennero apportate diverse modifiche rispetto al suo debutto negli anni '80. Oltre che alla linguistica descrittiva, esso attinse anche ad altre discipline come la logica, l'informatica e la statistica affinché i criteri che lo caratterizzavano potessero essere rappresentati in una modalità utilizzabile dalla macchina. In questo modo la grammatica formale, definita come un insieme di regole che descrivono un set di elementi finiti capaci di rappresentare la sintassi di una data frase, e più in generale di una data lingua, divenne un rigoroso metodo matematico in cui il linguaggio era trattato come un codice esatto e puntuale. Per molti versi, tale approccio rappresentò un recupero delle idee originariamente espresse da Weaver negli anni '50. Gli strumenti di traduzione automatica che vennero sviluppati a partire da esso operavano una duplice analisi del testo, la prima finalizzata a identificare le diverse parti della frase, la seconda ad assegnare a ognuna di esse delle informazioni sulla funzione sintattica svolta all'interno della frase stessa (Thurmain, 2005: 99). Per tale motivo furono detti *rule-based systems* (RBMT), essendo basati sul riconoscimento del ruolo svolto da ciascuna parola nella formazione della frase. Anche denominati sistemi "knowledge-based" e successivamente identificati come sistemi "classical approach", si trattava di sistemi *language-dependent*, ovvero sviluppati per una specifica combinazione linguistica. La loro principale caratteristica era rappresentata dall'utilizzo di dizionari e grammatiche contenenti informazioni linguistiche, semantiche, morfologiche e sintattiche sia della lingua di partenza che di quella di arrivo. Tali informazioni venivano applicate ai testi da tradurre per comprenderne la struttura e ricavare dei modelli da riprodurre nella lingua di arrivo. Il processo è suddiviso in tre fasi, come suggerisce lo schema riportato da Hutchins e Somers (1992: 72)

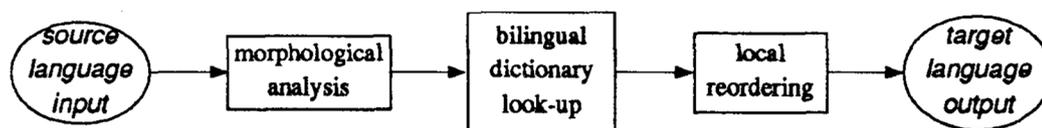


Fig. 26: I cinque passaggi dei sistemi RBMT (Hutchins e Somers, 1992)

Dato un input in una L1, il sistema procede a una sua analisi morfologica che attribuirà ad ogni elemento la corrispondente parte del discorso (articolo, nome, verbo ecc.), con ulteriori informazioni circa la flessione del verbo (forma, tempo verbale, persona) e la forma di nomi, aggettivi e articoli (singolare/plurale, maschile/femminile).

Successivamente a questo passaggio, si procederà alla traduzione di ogni parola tramite la ricerca dei corrispettivi nei dizionari di riferimento. Infine, tramite il ricorso alle grammatiche della lingua di arrivo, la frase verrà ricostruita e riordinata secondo le regole morfosintattiche della L2. Oltre alle grammatiche e ai dizionari delle due lingue, tra gli strumenti adottati per svolgere tali compiti vi sono un sistema di analisi morfologica della L1, un parser per l'analisi sintattica della L1, il vero e proprio traduttore automatico, un generatore morfologico per la L2 che lavora congiuntamente a un parser per comporre le frasi nella lingua di arrivo.

Come è facile comprendere, tale sistema può rivelarsi ottimale per frasi dalla struttura molto semplice, mentre per periodi più lunghi e complessi, composti da più frasi interconnesse da vincoli di coordinazione e subordinazione, richiederebbe una notevole capacità di analisi e il riconoscimento di strutture sintattiche articolate, difficilmente ottenibile dai sistemi elettronici. Inoltre, alla già difficile analisi e mappatura di periodi tanto complessi si aggiunge il problema della trasposizione nella L2. Se infatti è relativamente semplice trovare l'equivalenza di una frase elementare fatta di soggetto, verbo e complemento, quando si passa a strutture composite si devono fare i conti con la diversità di organizzazione dei periodi e i composti e le espressioni proprie nelle diverse lingue. Le collocazioni, ad esempio, o i termini e le espressioni composte, sono elementi particolarmente problematici per la traduzione automatica, a causa delle loro caratteristiche morfosintattiche, e per i vincoli di dipendenza che creano all'interno della frase, anche quando i termini che li compongono non si trovano in prossimità o nella loro posizione "canonica". Inoltre, in molti casi, per tali espressioni non esistono forme esattamente corrispondenti tra lingue diverse, anzi è molto facile tradurli in maniera erronea perché esse sono tipicamente composte da parole molto comuni, che unite formano tuttavia un composto semanticamente inequivocabile (Wehrli et al., 2009). Per far fronte a questo come ad altri problemi che necessitano di una profonda disambiguazione e della corretta identificazione degli elementi e delle loro funzioni nel testo, occorrerà fornire ai sistemi RBMT sufficienti strumenti e risorse a cui attingere, ma il rischio di erronei risultati è comunque reale per la complessità del processo di recupero e conversione di tali peculiari elementi. Proprio per far fronte a tali difficoltà, i modelli *rule-based* vengono spesso adoperati, con risultati interessanti dal punto di vista dell'accuratezza, nel tradurre i già citati linguaggi controllati. La loro grammatica

essenziale e lineare consente al programma il facile riconoscimento dei modelli e delle strutture delle frasi da tradurre, e la loro relativamente facile trasposizione nella lingua di arrivo (Mitamura e Nyberg, 1995; Allen e Hogan, 2000).

### 5.3.3. Gli approcci *transfer-based* e *interlingua-based*

A partire dalla metà degli anni '80, la categoria dei sistemi *rule-based* subì al suo interno un processo di differenziazione, man mano che nuovi modelli venivano creati, per tentare di restringere il numero di regole necessarie all'analisi testuale, e semplificare i processi coinvolti. In particolare, vennero distinti tre sistemi. L'approccio "diretto" o *dictionary-based* è tra questi la forma più elementare basata su una corrispondenza parola per parola, con o senza una preventiva analisi morfologica o lemmatizzazione delle forme flesse. Probabilmente i due tipi più interessanti di sistemi RBMT sono quelli basati sugli approcci così detti *transfer-based* e *interlingua-based*. Entrambi partono dall'ipotesi dell'esistenza di una rappresentazione intermedia tra la L1 e la L2, che catturi il "significato" della frase di partenza prima di trasferirlo tradotto nella lingua di arrivo. Per i primi, tale rappresentazione intermedia è dipendente dalla combinazione linguistica interessata, per i secondi invece essa non è legata alle lingue ma è una pura unità di concetto. In generale, i sistemi *transfer-based* applicano dei set di regole linguistiche definite come corrispondenze tra la struttura della lingua di partenza e quella di arrivo. L'analisi del testo di partenza dal punto di vista morfosintattico e semantico consente di creare una rappresentazione interna a partire dalla quale, utilizzando sia i dizionari bilingui che le grammatiche, verrà prodotta la traduzione. Tale rappresentazione è astratta e, seppur legata alla lingua di partenza, scevra da ogni informazione formale e ridotta all'essenza del contenuto. Il trasferimento avverrà prima da tale rappresentazione ad una pari rappresentazione nella lingua di arrivo. Solo dopo questo passaggio di trasferimento verrà data una forma linguistica alla rappresentazione concettuale. Nei sistemi *interlingua-based*, invece, tale rappresentazione, detta appunto interlingua, è libera da vincoli con le lingue sia di partenza che di arrivo. Per questo motivo, non è previsto il passaggio tra due rappresentazioni come nei sistemi *transfer-based*, ma la diretta produzione della traduzione a partire dall'interlingua stessa (Mitamura *et al.*, 1991). Questa strategia si è rivelata in grado di fornire risultati interessanti e con un grado relativamente elevato di accuratezza. Essa, al contrario dei modelli RBMT classici, offre il vantaggio di poter utilizzare gli stessi moduli di analisi di una L1 per tutte le altre lingue

di arrivo possibili, invece che per un'unica combinazione linguistica (Hutchins e Somers, 1992: 74). Il sistema interlingua presentato da Wilks nel 1972 per tradurre testi dall'inglese al francese è uno dei primi ad aver adottato questo modello. Esso non fa ricorso ad alcuna grammatica generativa, ma desume per ogni frase del testo originale le principali relazioni e strutture semantiche che andranno a formare la rappresentazione – interlingua – che verrà a sua volta successivamente trasposta nella lingua di arrivo tramite stereotipi, ovvero “*strings of French words, and functions evaluating to French words, which are attached to English word senses in the dictionary and built into the interlingual representation by the analysis routines*” (Wilks 1972: 6). Il programma di generazione riceve quindi una rappresentazione interlinguistica che contiene già il testo in francese e le procedure implicite che servono ad assemblarlo. Più di recente, Mitamura *et al.* (1991) hanno fuso i benefici del sistema *interlingua-based* con quelli offerti dall'utilizzo dei linguaggi controllati, che renderebbero indubbiamente più facile la creazione di una corretta rappresentazione nell'interlingua e la trasposizione di questa rappresentazione in una o più lingue di arrivo allo stesso tempo, riducendo inoltre in maniera drastica il bisogno di post-editing.

Sebbene tali sistemi mostrassero un notevole grado di sofisticazione rispetto ai primi modelli, e dessero adito a speranze di concreto miglioramento nel perseguimento degli obiettivi di una traduzione di buona qualità, anche essi non condussero, almeno nei primi anni, ad un sostanziale progresso in tal senso. Pur funzionando relativamente bene con testi lineari e controllati, essi erano incapaci di gestire in maniera ottimale testi comuni non appositamente preparati. La proposta di elaborazione delle regole grammaticali e sintattiche di una lingua secondo criteri e logiche matematiche era destinata a dare non pochi problemi, soprattutto nel tentativo di stabilire una corrispondenza perfetta tra i modelli e le strutture di due lingue al fine di trasporre il contenuto di un testo dall'una all'altra. Tuttavia, malgrado queste difficoltà, la sperimentazione sui modelli *rule-based* è proseguita fino ai giorni nostri, sviluppando sistemi e criteri operativi capaci di fornire migliori risultati rispetto agli inizi.

#### **5.3.4. I sistemi *corpus-based***

Dagli anni Novanta si apre la fase moderna della traduzione automatica. Le analisi sintattiche o semantiche vennero momentaneamente abbandonate a vantaggio del ricorso a database di testi bilingui allineati – prevalentemente corpora paralleli – nei quali

ricercare corrispondenze e analogie col testo da tradurre. In America, IBM applicò al sistema “Candide” un approccio statistico, mirante ad estrapolare in maniera automatica le traduzioni in base a calcoli statistici sulla frequenza di utilizzo di termini ed espressioni in date situazioni linguistiche, mentre in Giappone venne sviluppato il primo sistema “example-based”, capace di recuperare la traduzione dagli esempi di utilizzo dei linguaggi contenuti nei database o corpora di riferimento. Questi due nuovi approcci, statistico da un lato ed *example-based* dall’altro, vennero racchiusi all’interno della categoria *corpus-based MT*, data la loro comune dipendenza dall’utilizzo dei corpora, giustificata dal bisogno di accrescere la coerenza delle traduzioni automatiche grazie al ricorso ad esempi reali e concreti di utilizzo del linguaggio. Queste due nuove metodologie non scalzarono completamente le precedenti, che continuarono ad essere sperimentate in alcuni ambienti accademici. Venne così a crearsi un dualismo ad oggi tutt’ora persistente all’interno della disciplina, da un lato la ricerca sui sistemi *rule-based*, dall’altro quella sui sistemi *corpus-based*.

L’idea alla base dei modelli statistici ed *example-based* è identica, a cambiare sono le metodologie di estrazione delle informazioni ricercate (Rico Pérez e Sanchez, 1997), nel primo caso basate su calcoli probabilistici, nel secondo su modelli desunti da esempi pregressi. I corpora utilizzati all’interno di tali sistemi devono essere adatti alla loro applicazione in ambito linguistico. Ciò comporta la necessità di introdurre delle annotazioni, informazioni utili alla corretta interpretazione linguistica, e l’utilizzo di un’avanzata tecnica di allineamento che identifichi correttamente le corrispondenze frase per frase tra una lingua e l’altra.

I sistemi *example-based* mirano a produrre una traduzione automatica ricercando prima di tutto all’interno del corpus il segmento più simile alla frase nella lingua di partenza. Sarà quindi il grado di somiglianza tra la frase da tradurre e i segmenti analoghi nella lingua di partenza a determinare il recupero, in un passaggio successivo, della corrispettiva traduzione (Lopez, 2008). Tale apparentemente semplice processo è in realtà diviso in diversi passaggi, ognuno dei quali necessita di una specifica programmazione. Prima di tutto, occorre creare un modello traduttivo che stabilisca gli schemi da seguire. Tale modello è composto da equivalenze contestuali tra la lingua di partenza e quella di arrivo, che vengono estratte direttamente dal corpus parallelo e codificano in maniera esplicita o implicita tutte le informazioni linguistiche necessarie per la traduzione.

Costruire tale modello implica essenzialmente due parti, l'estrazione delle informazioni e la formalizzazione di tali informazioni entro degli schemi. Sono attività che condividono molto con le sperimentazioni sulle intelligenze artificiali e l'elaborazione del linguaggio naturale (Calzolari e Lenci, 2004), essendo legate fundamentalmente alla codifica delle conoscenze reali e concrete provenienti dal mondo reale. L'estrazione delle informazioni può avvenire manualmente, da parte di un gruppo di esperti che analizzano il corpus parallelo e decidono il tipo di regole sulla base delle quali creare dei template, o in maniera semi-automatica, grazie all'adozione di algoritmi e tecniche di intelligenza artificiale per l'identificazione e corrispondenza delle caratteristiche semantiche nelle due lingue. È importante sottolineare il carattere contestuale degli equivalenti traduttivi, poiché è questa qualità che consente di trasporre le informazioni relative alla lingua di partenza nelle corrispondenti nella lingua di arrivo. Il modello traduttivo estratto dal corpus bilingue stabilisce delle equivalenze dirette tra lingua di partenza e di arrivo all'interno di un contesto, prevenendo in questo modo associazioni erronee private di ogni riferimento contestuale. Una volta codificato il modello traduttivo, si procede con la ricerca del miglior match disponibile tra il testo da tradurre e tale modello. Questo processo, generalmente noto come "calcolo della distanza", implica il calcolo del grado di similarità tra lo schema di partenza e il testo da tradurre (Rico Pérez e Sanchez, 1997). Il livello di similarità viene spesso calcolato parola per parola, sulla base del paradigma morfologico di ciascuna di esse o a livello di sintassi dell'intera frase. Una volta calcolata la distanza, il sistema selezionerà la regola o il modello più appropriato.

### **5.3.5. I sistemi *example-based***

L'idea di applicare il metodo statistico alla traduzione automatica si afferma alla fine degli anni '80. Il suo scopo è quello di calcolare la probabilità di un possibile testo target di essere la traduzione di un testo di partenza, e adottarlo come base per ulteriori traduzioni. Non si tratta quindi di imitare, ma di tradurre a partire dalla stima delle probabilità che, date due frasi nelle lingue di partenza e di arrivo, una sia la traduzione dell'altra (Brown, 1990). Mentre nei sistemi *example-based* i modelli della TA erano rappresentati da condizioni e variabili applicabili ai testi di partenza e di arrivo, nei sistemi statistici questi modelli sono invece di tipo probabilistico. I parametri di tali modelli vengono determinati in maniera automatica sulla base di un ampio database di frasi allineate nelle due lingue, utilizzando due modelli probabilistici: un modello del

linguaggio e un modello della traduzione. Il modello del linguaggio stabilisce la probabilità di occorrenza di una singola parola sulla base delle parole che la precedono o la seguono nella frase, in modo da stabilire possibili legami tra parole, come nel caso in cui esse formino una collocazione o un'espressione composta, ma anche semplicemente per comprendere il contesto in cui dati termini vengono utilizzati. Il modello della traduzione calcola invece le probabilità che una parola nella lingua di partenza produca almeno una parola nella lingua di arrivo, poiché non tutte le parole di una frase di partenza possono avere un esatto corrispettivo nella frase di arrivo. Occorre quindi introdurre una misura di "fertilità" (Rico Pérez e Sanchez, 1997: 613) per stabilire la probabilità che una data parola produca zero, uno o più nella lingua di arrivo. Dopo aver determinato i modelli, le probabilità di traduzione vengono calcolate convertendo le parole e le espressioni della lingua di partenza nelle corrispondenti espressioni nella lingua di arrivo tramite il calcolo della probabilità che una parola della lingua di partenza corrisponda ad una o più parole nella lingua di arrivo (Barrachina *et al.*, 2009).

Tale approccio richiama il modo di procedere proprio dei CAT Tools, che ricercano nelle memorie di traduzione porzioni di testo già tradotte, ne valutano in termini percentuali la pertinenza rispetto all'input iniziale, e le sottopongono al traduttore affinché le convalidi o le modifichi a seconda dei casi. Non a caso lo sviluppo dell'approccio statistico alla traduzione automatica è avvenuto di pari passo con quello dei primi strumenti di traduzione assistita, con l'applicazione della linguistica dei corpora alla traduzione e lo sviluppo delle memorie di traduzione prodotte tramite il processo dell'allineamento di testi frase per frase tipico dei corpora bilingui. Vi sono quindi significative convergenze tra questi due strumenti<sup>40</sup> (Hutchins, 2001: 11).

Questi nuovi metodi di traduzione sono per certi versi più vicini al processo umano di traduzione, in quanto il cervello umano non traduce parole, bensì contesti e situazioni. Nell'estrarre le equivalenze traduttive dal corpus, le frasi verranno presentate all'interno del loro contesto di utilizzo, cosa che ne agevolerà la comprensione e consentirà di valutare se la traduzione proposta è adeguata al nuovo contesto traduttivo.

I modelli statistici si sono evoluti molto dal loro esordio ad oggi, potendo contare su di una sempre maggiore disponibilità di dati da cui attingere che hanno contribuito a migliorare anche la qualità dell'input. Inoltre, tali modelli si sono ben prestati negli anni

---

<sup>40</sup> L'aspetto verrà approfondito nel corso del prossimo capitolo (cf. 6.2).

a numerose sperimentazioni che mirano al miglioramento dei risultati tramite la formulazione di nuove modalità di analisi o di allineamento, o fornendo al programma informazioni sintattiche o semantiche utili a distinguere durante il recupero degli equivalenti su base probabilistica quelle parole che possono essere causa di ambiguità e fraintendimento (cf. 6.1). Si è sperimentato ad esempio un doppio allineamento, sia a livello di frasi che a livello di parola (Och e Ney, 2004), che sembra consentire una traduzione più corretta della frase (intesa non nel suo significato sintattico, ma come insieme di parole) attraverso l'adozione di un modello log-lineare. L'allineamento a livello di parola sfrutta le annotazioni aggiunte ai database di riferimento, che forniscono informazioni sulla categoria di appartenenza della parola (tanto grammaticale quanto tematica) e informazioni contestuali utili a tradurre le frasi in maniera più coerente. Questo approccio è ad oggi uno dei più promettente, perché i modelli sviluppati in base ad esso sono semplici da programmare e capaci di identificare e risolvere in diversi casi i problemi di riordino nella frase di arrivo, la traduzione di espressioni idiomatiche o di sottostringhe di testo come frasi o espressioni incidentali. Un ulteriore contributo per la soluzione di questi problemi è dato dall'adozione di un modello gerarchico (Chiang, 2007) che riesca a tenere in considerazione sia gli elementi interni alla frase (parole ed espressioni o *sub-phrase* come le chiama l'autore) sia porzioni più ampie di testo, e grazie al confronto tra questi diversi livelli testuali "apprendere" il corretto ordine delle parole nella lingua di arrivo. Le informazioni grammaticali e sintattiche in questo modello non provengono dall'annotazione del database, ma sono indotte dal testo parallelo. Le frasi allineate e ordinate in maniera gerarchica possono essere acquisite dal programma senza il bisogno di ulteriori informazioni sintattiche, e sembrano migliorare in maniera significativa l'accuratezza della traduzione (*ibid.*).

## 5.4 Ostacoli e criticità

### 5.4.1. Risultati parziali

L'exkursus compiuto in questo capitolo, non esaustivo ma comunque sufficientemente ampio da comprendere i principali modelli e sistemi sviluppati dall'inizio del fenomeno sino ad oggi, si conclude con l'ammissione dell'ancora insormontabile imperfezione dei sistemi di traduzione automatica. Malgrado i progressi

compiuti, e i risultati promettenti ottenuti soprattutto nel campo dei linguaggi controllati, emergono sempre delle criticità nel raggiungimento di una traduzione automatica di qualità. La TA si è dimostrata utile da un punto di vista economico, ma solo laddove l'input sia controllato, la terminologia standardizzata e la documentazione ripetitiva, e qualora venga utilizzato un linguaggio ristretto. Per le tipologie testuali più complesse, tuttavia, l'applicazione della TA è ancora - e molto probabilmente sarà sempre - incapace di fornire risultati accettabili in termini di qualità.

#### **5.4.2. Individuare le criticità**

Perché i sistemi di traduzione automatica non raggiungono le aspettative? Il primo ostacolo sta nella grande complessità e nella limitata prevedibilità dei linguaggi naturali (Och e Ney, 2004: 417): numerose parole hanno più significati e diverse possibili traduzioni; le frasi possono essere interpretate in diversi modi e spesso le relazioni tra le parti del discorso sono vaghe e ambigue. L'uomo non percepisce questa complessità in maniera problematica, perché il cervello mette in moto in maniera automatica i meccanismi di comprensione e produzione, e può fare inoltre ricorso alle conoscenze del mondo per risolvere le varie ambiguità, cosa di cui il computer non è in grado. Proprio la vivacità delle lingue fa della tradurre un'attività per certi versi *creativa*, poiché in molti casi il traduttore deve "inventare" soluzioni a problemi traduttivi quali la mancanza di un preciso equivalente traduttivo, la presenza di neologismi nel testo di partenza, il bisogno di agire da "mediatori culturali" tra il testo di partenza e l'audience che lo fruirà in traduzione. I computer sono fondamentalmente dei dispositivi che seguono delle regole chiare, univoche e lineari (Arnold, 2000: 2), e di fronte alla creatività linguistica non riescono a fornire una risposta adeguata.

Per Hutchins (2005: 21) le ragioni del successo parziale della traduzione automatica sono le stesse che portano un traduttore a produrre un testo di scarsa qualità, sintetizzabili in un'insufficiente conoscenza del tema trattato e della lingua di partenza o di arrivo. Il computer può arrivare ad avere una sufficiente conoscenza solo di linguaggi controllati e ridotti all'essenziale, e solo di argomenti molto precisi che utilizzino un vocabolario circoscritto ad un numero limitato di termini. Tuttavia, anche in questi casi il rischio di errore è molto probabile, perché il linguaggio, anche nelle sue espressioni più semplici, è inscindibilmente legato alla cultura, elemento che i computer non sono in grado di cogliere. Inoltre occorre tenere presente che la traduzione di per sé è un'attività difficile

e complessa, che richiede diverse abilità e che spesso non si esaurisce alla semplice trasposizione di un contenuto da una lingua a un'altra. Il testo prodotto deve riprodurre lo stile dell'originale, essere quindi elegante, poetico, esplicativo, persuasivo, interessante a seconda del caso (Arnold, 2000: 1). La traduzione non è mai fine a se stessa, ma è mossa da un bisogno, un'utilità finale. Mentre per l'uomo comprendere l'intento comunicativo di un testo è molto semplice, e il bisogno di riprodurlo al meglio in traduzione un'ovvia necessità, per i software di traduzione è pressoché impossibile fare distinzioni di stile, forma e funzione del testo. Per Arnold (2000)

*“At the root of these problems are four particular limitations of computers, namely, the inability of computers to: (i) perform vaguely specified tasks; (ii) learn things (as opposed to being told them); (iii) perform common sense reasoning; (iv) deal with some problems where there is a large number of potential solutions.”*

Oltre a questa sintetica descrizione, l'autore scende nel dettaglio analizzando il processo di traduzione automatica e individuando i principali ostacoli e le possibili difficoltà insorgenti nei sistemi *transfer-based* e in quelli statistici, i due principali approcci tanto al momento in cui l'autore ha scritto quanto oggi. Relativamente ai sistemi *transfer-based*, Arnold identifica dei “problemi di analisi”, laddove il software non è in grado di comprendere correttamente il testo di partenza; dei “problemi di sintesi” che riguardano l'erronea espressione di un particolare contenuto; dei “problemi di trasferimento” quando si presentino irriducibili differenze tra i modi in cui uno stesso contenuto è espresso in lingue diverse; infine, dei “problemi di descrizione” laddove vi sia un errore nell'insieme delle risorse raccolte per costruire il sistema di TA. In modi diversi, queste criticità sono presenti in entrambi i modelli di TA, e ne determinano la qualità dei risultati. La traduzione automatica non è quindi in grado di produrre un testo che presenti una forma adeguata per la lingua di arrivo e preservi al contempo il contenuto del testo di partenza, perché forma e contenuto si determinano a vicenda, ed esistono differenze significative nel modo in cui lingue diverse esprimono lo stesso contenuto. Queste differenze non sono inoltre facilmente codificabili e rappresentabili con comandi precisi, né univocamente identificabili in un approccio statistico (Arnold, 2000: 20).

Per ovviare alle inevitabili imprecisioni o errori contenuti nel testo tradotto automaticamente, laddove ciò sia necessario per gli scopi dell'utente finale, in particolare come si è detto quando il testo tradotto debba garantire determinati standard qualitativi, è necessario procedere con un post-editing svolto da un traduttore (Hutchins 2005: 6; 2001:9). Ciò può rappresentare un ulteriore ostacolo in quanto prevede dei costi che in genere i fruitori della TA intendono evitare. Laddove il grado di post-editing necessario sia molto elevato, il ricorso alla traduzione automatica diventerebbe poco conveniente o interessante. Gli errori più importanti riportati in genere in fase di revisione sono causati dalla difficoltà del computer di gestire frasi complesse e articolate, come periodi lunghi con diverse preposizioni subordinate. Altri errori possono essere causati da problemi legati al testo di partenza, come refusi o punteggiatura carente. In altri casi si tratta di errori di traduzione veri e propri, laddove il computer compia un'associazione di termini impropria o inadatta al contesto della frase, come spesso accade. La quantità di post-editing necessaria varia in base al grado di standardizzazione del testo, in quanto come si è detto, più standardizzato e "prevedibile" è il testo, più sarà facile tradurlo automaticamente (Hutchins, 2005). Un maggior grado di creatività e libertà nell'espressione causerà al contrario maggiori difficoltà al software e necessiterà quindi di un più incisivo post-editing. Inoltre, maggiore sarà stata la quantità e qualità di pre-editing, finalizzata proprio ad evitare preventivamente i problemi nella traduzione, minore sarà il bisogno di post-editing (Hutchins 2005: 23), ma ciò comporterà comunque un intervento umano e dei costi, che rappresentano spesso il fattore discriminante per agenzie e privati nello scegliere se avvalersi o meno del servizio di traduzione automatica al posto di una traduzione umana svolta con sussidi diversi.

#### **5.4.3. I limiti materiali**

Un ulteriore aspetto critico da valutare riguarda la fase di implementazione dei sistemi, che richiede un lungo e complesso lavoro di raccolta dei materiali e delle risorse di riferimento, l'adozione delle grammatiche e delle regole di analisi e trasferimento, la definizione degli algoritmi di ricerca etc. Questo processo deve essere compiuto per ogni lingua e per ogni combinazione linguistica, tanto per i sistemi *rule-based* quanto per quelli *example-based*, e implica dei costi non indifferenti e un lavoro molto articolato e impegnativo. Infatti, le regole e gli algoritmi che si applicano per un sistema di traduzione inglese – francese non potranno applicarsi anche per sistemi inglese – russo o inglese –

giapponese; le relazioni tra le lingue sono diverse, le risorse disponibili altrettanto, la programmazione di ogni sistema è un lavoro che accomuna esperti di diversi background professionali, e molte società di servizi linguistici o di IT valutano bene costi e profitti prima di decidere se implementare o meno sistemi specifici, soprattutto per quelle lingue di minore rilevanza economica o strategica. Per questo motivo, fino all'insorgere dei sistemi gratuiti on-line di cui si parlerà al capitolo successivo, la maggior parte dei software di TA era disponibile esclusivamente per limitate combinazioni linguistiche (Dillinger e Lommel, 2004: 6), quelle che avrebbero certamente garantito un buon ritorno sugli investimenti.

Persino dopo l'implementazione dei sistemi occorre tenere presente il bisogno di un costante aggiornamento dei software, per stare al passo con le rapide evoluzioni delle lingue e soprattutto dei linguaggi specialistici. Un'architettura tanto complessa come quella sottostante ai sistemi di TA non è facilmente modificabile; la semplificazione dei meccanismi di aggiornamento è una delle principali attenzioni di studiosi e sviluppatori, in vista delle esigenti richieste del mercato. Sono state sperimentate delle modalità di correzione ed accrescimento delle risorse in tempo reale durante la traduzione (Thicke, 2013), più fattibile per i sistemi statistici sull'esempio della modalità di aggiornamento delle TM nei CAT Tools. D'altro canto, i sostenitori dei modelli *rule-based* sostengono che essendo le grammatiche solitamente più stabili della terminologia, tali sistemi avranno solamente bisogno di un aggiornamento costante dei glossari di riferimento (Coughlin, 1990).

Un modo per tentare di superare gli ostacoli e le criticità dei sistemi di TA è rappresentato dal tentativo di sfruttare congiuntamente i benefici dei diversi modelli esistenti, all'interno di sistemi detti "ibridi", come si vedrà di seguito, al capitolo sesto.

## Capitolo 6: La traduzione automatica oggi

### 6.1 *Modelli classici e ibridi*

#### 6.1.1. **Un'offerta differenziata**

Passando a discutere in questo capitolo della situazione della traduzione automatica oggi, non si può che partire affermando che questo fenomeno si è diffuso in maniera capillare, con la moltiplicazione di strumenti e servizi offerti in particolare sul luogo degli scambi e della comunicazione per eccellenza che è il Web. Esistono servizi gratuiti online e software a pagamento sia per gli utenti comuni che per i traduttori. Destinati a questi ultimi, riconosciuti finalmente come potenziali utilizzatori di tale servizio, sono in particolare dei sistemi specializzati, settoriali o limitati ad una precisa combinazione linguistica, sviluppati appositamente per tentare di fornire un contributo realmente utilizzabile dai traduttori e che ne accresca la produttività (Champsaur, 2013: 21). Anche le case produttrici dei CAT Tools si sono munite di loro personali sistemi di TA, che vengono proposti come alternativa alle TM o unitamente ad esse per un uso congiunto (cf. 6.2). Ad esempio, la società SDL ha lanciato il proprio portale di traduzione automatica a pagamento (SDL BeGlobal<sup>41</sup>), con l'intento di offrire un servizio principalmente per le aziende e le società che intrattengono scambi frequenti con clienti in lingue diverse, e che possono quindi trarre vantaggi da una traduzione automatica online in tempo reale che garantisca una maggiore garanzia di attendibilità rispetto ai servizi gratuiti, e che possa inoltre essere personalizzato sulla base delle necessità dell'utente, in particolare con l'adozione della terminologia specifica.

Ad oggi convivono sul mercato sistemi diversi, afferenti alle due principali tipologie sviluppate, sistemi statistici quali Google Translate, Moses, Asia Online, Bing, Language Weaver, e sistemi *rule-based* (RB) quali Systran, Lucy, Reverso, PROMT (Thicke 2013). La lista è tutt'altro che esaustiva, l'offerta sul mercato è ampia e rispondente, come si è

---

<sup>41</sup> <http://www.sdl.com/products/sdl-beglobal/>

visto, ai bisogni di un'utenza più o meno specializzata. Mentre gli utenti casuali optano per le soluzioni gratuite e immediate per risolvere i problemi sporadici di comprensione che non necessitano nella maggior parte dei casi di una qualità elevata, coloro che invece hanno esigenze speciali, come appunto le aziende, i gestori di pagine web e i prestatori di servizi linguistici come agenzie e traduttori, potranno scegliere tra le varie alternative lo strumento che possa tornare a loro favore in considerazione delle specifiche esigenze di lavoro (Coughlin, 1990). Oltre alla scelta di base tra CAT Tools e TA, o l'uso congiunto di entrambi, occorre sapere individuare il miglior CAT Tool o il migliore sistema di TA sulla base di alcune considerazioni. Nella fattispecie, quando si deve scegliere tra i diversi modelli di TA, occorrerà tenere conto di aspetti specifici e tecnici. Occorre verificare che un dato sistema supporti il contenuto, il formato file e la combinazione linguistica del caso, e che sia adatto alla tipologia testuale da tradurre. Ad esempio, i sistemi RB sono in genere più adatti a contenuti molto specifici e circoscritti, mentre contenuti più ampi e generici sono meglio elaborati dai sistemi statistici (Thicke, 2013: 12). Alcune lingue, come il giapponese e il tedesco, sembrano essere più adatte ad un sistema RB, mentre le lingue meno comuni possono avvalersi solo di sistemi statistici, non essendo stati sviluppati specifici sistemi RB per loro (cf. 7.4). Se si hanno a disposizione numerose risorse da utilizzare nel corso del lavoro, come delle TM o documenti allineati, allora la TA statistica garantirà una migliore performance. In mancanza di risorse abbondanti, i sistemi RB, operanti sulla base di un set di regole grammaticali e della terminologia specifica saranno i più idonei (*ibid.*). Altre considerazioni possono riguardare il modo in cui si gestirà il lavoro, se in maniera condivisa con altri colleghi o in autonomia, o se si pensa di avvalersi di un sistema personalizzato o generico, o ancora se il formato del file richiede la protezione dei tag per la formattazione, ecc.

### **6.1.2. Sistemi personalizzati**

Un altro elemento da considerare è la necessità di personalizzare e preparare il sistema di TA con la propria terminologia, dizionari e materiali di riferimento. Non tutti i sistemi lo consentono, spesso è utile ricorrere a sistemi specifici a pagamento. Inoltre, mentre i sistemi SMT sono *language-independent* come le TM, e possono quindi essere personalizzati potenzialmente per tutte le combinazioni linguistiche semplicemente fornendo un modello linguistico rappresentato da adeguati materiali di consultazione (corpora, documenti allineati, dizionari), lo stesso non si può dire dei sistemi RB, creati

per combinazioni linguistiche specifiche e basati sul riconoscimento di regole grammaticali e sintattiche tipiche delle lingue coinvolte. D'altro canto, tuttavia, laddove le combinazioni linguistiche siano presenti, occorrerà fornire solo la terminologia specifica e il sistema sarà già in grado di produrre risultati senza il bisogno di altra preparazione. In genere i sistemi RB sono preferiti per quelle traduzioni che richiedono un uso accorto e specifico della terminologia, essendo tali sistemi più facili da corredare della corretta terminologia, nonché per il maggiore controllo che essi rendono possibile. I sistemi statistici, necessitando di una più ampia documentazione e di maggiori risorse per poter operare, ed estrapolando la terminologia sulla base della sua frequenza di ricorrenza all'interno dei database di riferimento, non possono garantire la stessa precisione e sono quindi preferibili per altre tipologie testuali.

Indubbiamente i sistemi personalizzati sono quelli in grado di fornire i risultati migliori, anche se la loro programmazione è costosa e potrebbe far valutare l'utilizzo di tecnologie alternative, nella fattispecie le TM che nascono già per loro natura settoriali e specializzate. La personalizzazione comporta non solo l'adozione di una specifica terminologia di cui il cliente deve essere preventivamente in possesso, ma anche il ricorso a database settoriali ed annotati con informazioni utili ai fini dell'azienda o ente che la richieda ed equivalenti traduttivi autorizzati da questi (Hutchins, 2003: 11), ma ancora prima una predisposizione del sistema ad elaborare – laddove necessario – particolari tipologie testuali o formati file, l'utilizzo di un linguaggio di *mark-up* o altre caratteristiche tecniche, oltre che linguistiche. Si è stimato che un sistema di TA correttamente preparato e personalizzato possa produrre risultati notevolmente migliori (Thicke, 2013: 15). Prendendo come metro di valutazione la velocità di post-editing successivo alla traduzione, si è notato che la velocità di revisione dei risultati prodotti da un sistema di TA personalizzato e corredato delle necessarie risorse può quasi raddoppiare rispetto alla velocità di revisione per sistemi non personalizzati (da 4000 parole al giorno a 7400).

### **6.1.3. I modelli ibridi**

Un fenomeno a cui si è accennato alla fine del capitolo precedente è l'ideazione e sviluppo di modelli ibridi, termine con il quale si intendono principalmente quei sistemi che mettono insieme gli approcci *rule-based* ed *example-based*, al fine di ovviare ai limiti dei diversi modelli precedenti unendo i vantaggi da essi offerti affinché si sostengano a

vicenda per offrire un risultato di qualità più elevata (Thurmain, 2005). Si considerano inoltre ibridi quei sistemi che richiedono l'intervento umano in corso d'opera e quelli in cui traduzione automatica e CAT Tools vengono usati congiuntamente, come si vedrà al seguente paragrafo 6.2.

Prendendo come metri di valutazione tre parametri quali la correttezza grammaticale, la comprensibilità e la presenza di errori, Thurmain (2005) ha svolto un esperimento per valutare i punti di forza dei sistemi RB e statistici. I risultati hanno dimostrato che, sebbene i sistemi RB forniscano una migliore correttezza grammaticale e di conseguenza una minore presenza di errori, la traduzione compiuta tramite il sistema statistico risulta generalmente più comprensibile, e le frasi prodotte non sono né perfette né eccessivamente errate. Poiché la traduzione automatica di tipo statistico si basa sulla consultazione di testi allineati le cui porzioni (chunk) sono correlate secondo criteri e algoritmi che ne permettono il recupero all'occorrenza, laddove la corrispondenza di tali chunk con un nuovo segmento da tradurre sarà alta, la qualità della traduzione sarà buona; laddove vi sia un livello inferiore di corrispondenza, i risultati saranno errati o imperfetti. Uno dei problemi principali per i sistemi statistici è la mancanza di quella componente grammaticale e sintattica che possa prevenire i frequenti errori di flessione di nomi e aggettivi (soprattutto traducendo da una lingua come l'inglese, che ha un'unica forma di singolare e una di plurale, non flessa in base al genere), di erronea identificazione delle parti del discorso e il mancato riordino delle parole all'interno della frase prodotta, causato proprio dalla mancanza di un modello sintattico di riferimento. I sistemi *rule-based*, d'altro canto, nel loro tentativo di suddividere il testo di partenza per analizzarne la sintassi e riprodurlo sulla base dei dizionari e delle grammatiche della lingua di arrivo, corrono il rischio di compiere una traduzione erronea quando le strutture sintattiche di due lingue sono molto diverse, e al contempo la loro dipendenza dai dizionari senza la determinazione del giusto contesto di utilizzo può far sì che la terminologia venga utilizzata in maniera impropria. Si può quindi dire che i punti di forza per i modelli RB e per quelli statistici siano da un lato l'analisi sintattica e grammaticale, dall'altro la presenza di contesti di utilizzo reale. Unificando tali aspetti in un unico sistema, si otterrebbe una traduzione automatica dalle prestazioni migliorate. I due sistemi affrontano la questione da due prospettive completamente diverse, per comprendere le quali si può prendere in prestito la dicotomia tra competenza ed esecuzione puntualizzata

da Chomsky (1957). La prima identifica la capacità di un parlante di creare nuove frasi grazie all'applicazione di regole; la seconda indica invece l'utilizzo reale del linguaggio fatto di scelte e preferenze dal carattere soggettivo e peculiare. Mentre i sistemi RB, all'interno della tradizione della grammatica generativa, si concentrano sul primo aspetto, i sistemi statistici tendono maggiormente al secondo. Tuttavia, la linguistica insiste sull'importanza di entrambi gli aspetti e della loro complementarità per una piena e reale competenza linguistica. Anche nei sistemi che tentano di riprodurre i meccanismi dei linguaggi naturali è sempre più riconosciuta la necessità di far convergere i diversi approcci sperimentali in un modello linguistico più completo ed efficiente.

#### **6.1.4. Architetture possibili**

Come progettare un sistema di TA che integri i migliori elementi degli approcci esistenti all'interno di un'architettura ibrida? Il processo può prevedere l'arricchimento di un sistema statistico con risorse provenienti da un sistema *rule-based* o al contrario l'applicazione in un sistema RB delle funzioni tipiche dei modelli statistici. Il primo caso di ibridazione, vale a dire l'introduzione di informazioni linguistiche nei sistemi statistici, è il più sperimentato e si è dimostrato molto promettente per i risultati che ha ottenuto nelle diverse sperimentazioni (Bazrafshan e Gildea 2013; Riesa, Irvine e Marcu, 2011; Wang *et al.* 2010). Esso può essere implementato in due modi, o fornendo maggiori informazioni all'elaborazione statistica, o svolgendo un'analisi valutativa successivamente ad essa (Thurmair, 2005: 103). Nel primo caso, sarà necessario creare dei migliori dati in ingresso, annotandoli con le necessarie informazioni grammaticali e sintattiche che possano rendere possibile il riconoscimento di modelli e strutture sintattiche e contribuire a risolvere più facilmente problemi quali la difficoltà di risalire alle forme base dei verbi o all'analisi dei nomi composti, per una più corretta correlazione con i loro corrispettivi nella lingua di arrivo. Il secondo metodo, meno sperimentato, suggerisce l'introduzione di una seconda fase di analisi, questa volta non dei dati in ingresso ma delle traduzioni prodotte dal sistema, al fine di valutarle dal punto di vista linguistico e confermare la funzione sintattica esercitata da ogni componente. Le annotazioni che verranno allegate potranno essere utilizzate in traduzioni future per migliorare il risultato prodotto (*ibid.*).

Passando al caso opposto, l'introduzione di informazioni di tipo probabilistico nei sistemi RB implica unire all'analisi di tipo grammaticale e sintattico una distribuzione di pesi

diversi alle varie regole che possono applicarsi, sulla base della loro frequenza di utilizzo e di altri dati ricavati da un calcolo statistico. Nella fase di analisi iniziale, quindi, alle regole che vengono applicate per stabilire il modello sintattico e grammaticale della frase, si aggiungeranno delle informazioni relative alla frequenza di utilizzo osservata e altri dati statistici sul testo di partenza, come la possibile presenza di collocazioni ecc. Anche la selezione delle espressioni equivalenti nella lingua di arrivo non si baserà solo sul trasferimento del modello semantico-sintattico desunto dalla lingua di partenza, ma terrà anche conto degli aspetti legati alla esecuzione nella lingua di arrivo, sulla base di un'analisi probabilistica di tipici pattern linguistici corrispondenti alla lingua di uso reale. In particolare ciò consentirà di valutare parametri quali la selezione dei termini, l'identificazione di eventuali collocazioni nel testo di arrivo e il grado di prossimità del testo tradotto con le espressioni realmente utilizzate dai parlanti. Ciò avviene tramite strumenti appositi (come ad esempio la distanza di Levenshtein, cf. prossimo paragrafo 5.2). Questa modalità di ibridazione tra sistemi è stata largamente sperimentata, anche se ad oggi sono stati prodotti solo sistemi parziali, trattandosi di un processo molto complesso.

Come più volte accennato nel corso di questo capitolo, la qualità dei risultati prodotti dai diversi sistemi di TA è una variabile che nel corso degli anni è migliorata molto lentamente e in maniera non rispondente alle aspettative di molti (Hutchis, 2005). La ricerca sulle architetture ibride per i sistemi di traduzione automatica è oggi al centro dell'attenzione di numerosi studiosi per le potenzialità che presenta, proprio relativamente all'accrescimento della qualità dei sistemi di traduzione automatica. Sarà necessario risolvere problemi di tipo tecnico e condurre sperimentazioni a più ampio raggio, ma i presupposti promettenti offerti dagli studi già svolti (Thurmair, 2005) sono di incoraggiamento per coloro che intendano perseguire la finalità di una traduzione automatica di qualità. Una considerazione interessante che è possibile fare, (e che verrà ampliata al capitolo successivo, cf. 7.3) riguarda il confronto tra TA e traduzione umana. La prima è stata sviluppata nel tentativo di imitare e riprodurre attraverso il computer i complessi meccanismi cognitivi, linguistici e decisionali della seconda. Unire i punti di forza dei diversi modelli, mettendo insieme le conoscenze formali dei sistemi RB e la conoscenza dell'uso reale della lingua e della terminologia desumibile dalle risorse dei sistemi statistici potrebbe consentire di riprodurre in maniera più completa i processi

naturali legati alla traduzione. Essi si fondano su conoscenze di tipo linguistico ma anche pratico e generico, e a valutazioni che non sono solo dettate da regole precise, ma anche dal buon senso e dalla prassi quotidiana. I sistemi ibridi appaiono quindi come i migliori per tentare di riprodurre, nei limiti comunque presentati dai cervelli elettronici, i meccanismi umani legati alla traduzione. A controprova di ciò, i migliori risultati ottenuti in termini di qualità e naturalezza del testo tradotto.

## 6.2 Traduzione automatica e CAT Tools

Veniamo adesso ad una seconda forma di ibridazione, che si è molto diffusa nell'ultimo decennio circa, ovvero la tendenza ad un uso congiunto dei CAT Tools, fondamentalmente sistemi di sussidio alla traduzione svolta dall'uomo, con i sistemi di traduzione automatica, che potenzialmente tendono ad escludere il contributo umano se non in fase di pre- o post-editing (Hutchins 2005: 14; Bowker, 2002: 4). Sempre più CAT Tools ospitano nel loro ambiente di lavoro un sistema di traduzione automatica sviluppato dalla stessa casa produttrice, o un collegamento ad uno o più sistemi tra quelli disponibili sul Web. Allo stesso tempo, i sistemi di TA stanno sempre più incamerando le TM come componenti di sistema, e si avvalgono degli stessi algoritmi di recupero dei match utilizzati dai CAT Tools. La motivazione principale di tale convergenza è innanzitutto la possibilità di sfruttare congiuntamente i benefici di entrambi i sistemi e al contempo superare i limiti che essi presentano. Inoltre, i CAT Tools e la TA condividono numerosi strumenti e risorse, quali i sistemi di gestione della terminologia, i controlli di qualità, i metodi per l'acquisizione di nuova terminologia nel corso del lavoro, il ricorso a corpora paralleli e documenti allineati ecc. Una particolare somiglianza si può riscontrare tra le TM e i sistemi di TA *example-based* (cf. 4.3), ovvero quelli statistici o basati sui corpora, che propongono come i CAT Tools traduzioni desunte da testi allineati o corpora bilingui sulla base del grado di similarità col testo di origine. I sistemi *phrase-based* (cf. 4.2), invece, hanno in comune con i CAT Tools la divisione del testo in porzioni solitamente corrispondenti alla frase. Più di tutto, CAT Tools e TA hanno uno scopo comune, quello di contribuire a produrre traduzioni di qualità garantendo inoltre rapidità ed economicità del servizio, pur nella diversità delle loro metodologie.

### 6.2.1. Un contributo reciproco

Indubbiamente, la forma più comune e maggiormente diffusa di ibridazione tra i due sistemi è proprio l'introduzione della TA all'interno dei CAT Tools. Essa, in genere, viene utilizzata come pre-traduzione nei casi in cui non venga recuperato alcun match dalla TM. La sua capacità di fornire un output in tempo reale è molto comoda in tal senso, ma come si è più volte detto, i suoi risultati necessitano comunque di una revisione, specialmente quando le frasi tradotte dalla TA andranno a formare delle *translation units* salvate nella memoria di traduzione. Il rischio che si correrebbe in questo caso è che l'erronea traduzione venga riutilizzata anche in documenti successivi. Quindi, a meno di non migliorare significativamente il sistema di traduzione automatica, essa non può che essere relegata ad un compito secondario dall'utilità molto relativa, e spesso considerata dai traduttori come un'ultima spiaggia quando non si trovino risorse migliori e maggiormente affidabili.

Al di là della mera introduzione della TA all'interno dei CAT Tools, la cui applicazione è essenzialmente pratica e spesso ad esclusivo beneficio delle agenzie che risparmiano sul compenso del traduttore<sup>42</sup>, è interessante concentrarsi su quegli studi che propongono invece delle soluzioni per una profonda interrelazione tra i due sistemi, e comprendere come gli uni possano avvalersi beneficamente del contributo degli altri. I CAT Tools, per esempio, presentano vantaggi evidenti rispetto ai sistemi di TA, vantaggi fondamentalmente rappresentati dal grado di affidabilità delle traduzioni proposte, frutto da un lato di un accurato lavoro di creazione e manutenzione delle TM, e dall'altro da un più efficiente ed affidabile meccanismo di individuazione dei match. I due aspetti sono interconnessi, poiché più settoriale e ben mantenuta sarà la TM, più facile sarà estrarne risultati ottimali. In realtà, la principale differenza tra i due sistemi risiede indubbiamente nel diverso grado di partecipazione umana al processo traduttivo (Taravella e Villeneuve, 2013), che garantisce nel processo condotto con i CAT Tools un costante monitoraggio di ogni passaggio, e di conseguenza una maggiore qualità del prodotto finale. Le TM infatti non solo sono più accurate in partenza, ma lo divengono sempre più man mano che vanno arricchendosi in corso d'opera delle correzioni, aggiunte e modifiche apportate dal traduttore. Inoltre, la maggiore interattività offerta dai CAT Tools consente all'utente di

---

<sup>42</sup> Una frase tradotta automaticamente e poi revisionata dal traduttore viene solitamente conteggiata alla pari di un fuzzy match, quindi pagata circa la metà di una frase tradotta per intero.

effettuare ricerche mirate all'interno delle risorse disponibili, come la *Concordance search* o la consultazione manuale dei glossari. All'interno dei sistemi di TA, invece, l'interattività è volutamente ridotta al minimo, l'utente compie essenzialmente solo l'azione di inserire il testo da tradurre e selezionare la lingua di arrivo, il resto viene tutto svolto dal computer in automatico. Come verrà mostrato più in dettaglio nel prossimo paragrafo, alcuni sistemi di traduzione automatica online basati sull'utilizzo di TM (quindi sistemi ibridi) hanno da qualche tempo inserito una funzione che permette all'utente di contribuire alla produzione della traduzione, migliorando la soluzione ottenuta, che verrà salvata a beneficio degli utenti futuri. L'utente potrà anche visualizzare più possibilità di traduzione di uno stesso termine e decidere quale applicare. Si tratta di un modo semplice per introdurre nella traduzione automatica la componente umana e rendere possibile anche in essa un perfezionamento delle risorse, e rappresenta uno dei vantaggi più semplici ed immediati che l'adozione delle TM può offrire alla traduzione automatica. Tuttavia, il contributo maggiore che le TM possono fornire ai sistemi di TA è più profondo e riguarda le dinamiche di estrazione dei *best match* delle risorse utili.

Come suggerito da Simard e Isabelle (2009), il modo migliore per integrare sistemi TA e CAT è quello di sviluppare un sistema di traduzione automatica che sia il più possibile somigliante ai CAT tools, ovvero che si comporti come una TM. Ciò è possibile essenzialmente in due modi, ovvero fornendo alla TA la possibilità di estrapolare efficacemente i match con elevato grado di similarità, e agli utenti la possibilità di filtrare le traduzioni che abbiano maggiore probabilità di essere utili. Data la maggiore somiglianza dei sistemi *example-based* e *phrase based* con la TM, Simard e Isabelle hanno condotto uno studio al fine di sviluppare un sistema di TA di tale tipologia, capace di sfruttare i vantaggi degli *exact* e *fuzzy match* contenuti nella TM di un sistema CAT. La prima e più semplice possibilità per un'applicazione congiunta di TM e TA è quella di fornire ad entrambe gli stessi parametri per il calcolo dell'attinenza dei segmenti salvati rispetto ad una nuova query. Lo studio fa riferimento alla *distanza di Levenshtein*, o distanza di edit, che misura il grado di differenza tra due stringhe di caratteri. Tramite tale distanza può essere espressa la somiglianza di una nuova query (una nuova frase da tradurre) con le stringhe di riferimento (le TU presenti in memoria), ed è tale unità di misura a determinare la percentuale di match nelle memorie di traduzione. Minore sarà la distanza, maggiore sarà la somiglianza e quindi la percentuale di match. Oltre ad un dato

marginale di distanza, la TM non riconoscerà più la presenza di un match, e non proporrà quindi alcun suggerimento. Il saper determinare in maniera funzionale tale margine sarà un elemento importante nell'utilizzo ottimale di questa tecnologia. Applicando lo stesso criterio ad un sistema TM-TA, si potrà per esempio stabilire che sino ad un dato valore della distanza di Levenshtein verrà applicato il sistema TM, mentre per valori superiori, quindi per una minore corrispondenza tra TU e query, si applicherà invece la TA. Questo sistema, detto  *$\beta$ -combination* (Simard e Isabelle, 2009: 122), è una maniera estremamente semplice di far dialogare i due strumenti, ed offre la possibilità di determinare il valore di  $\beta$  a seconda delle preferenze dell'utente.

Un'altra possibilità, più complessa, è quella di programmare il sistema di TA affinché emuli i processi di estrazione e recupero dei match utilizzati dalle TM. In questo modo, la TA potrebbe essere in grado di sfruttare attivamente le risorse con un grado più elevato di corrispondenza ottenute dalla TM. A tal fine è necessario operare sul sistema di TA su molteplici fronti, dalla identificazione dei migliori materiali di riferimento alle capacità di decodifica e ricerca di soluzioni funzionali dalla TM. A tali processi viene applicata, oltre alla distanza di Levenshtein, anche la metrica WER (*word-error rate*), che valuta in percentuale il grado di utilità di una traduzione proposta sulla base della traduzione corretta presente in memoria, nonché diversi meccanismi per la stima della qualità (QE, *quality estimation*) basati su calcoli percentuali e probabilistici, relativi all'intera stringa di testo o a porzioni di essa, per valutare l'attinenza e la qualità della traduzione proposta anche come sommatoria di parti, ovvero dividendo la frase in porzioni minori per ognuna delle quali verranno calcolati parametri di similarità e attinenza con il contesto dell'intera frase. I meccanismi di QE consentono di filtrare quei risultati che hanno una minore probabilità di tornare utili al traduttore, facilitando l'identificazione delle risorse più correttamente applicabili alla situazione (Simard e Isabel, 2009).

### **6.2.2. Alcuni studi sull'argomento**

Un interessante studio è stato condotto da Biçici e Dymetman (2008), che hanno sviluppato una memoria di traduzione dinamica utilizzando un sistema statistico di TA allo scopo di migliorare il recupero di *fuzzy match* dalla TM. Gli studiosi infatti riconoscono nei *fuzzy match*, decisamente più frequenti degli *exact match*, il beneficio principale offerto dai CAT Tools, ma anche una parte considerevole del lavoro del traduttore e una delle principali voci di spesa per l'agenzia o il cliente. Da qui l'ipotesi di

automatizzare almeno in parte la revisione e correzione dei *fuzzy match* tramite l'introduzione di un sistema statistico di TA. Per tradurre una data frase *S*, il sistema recupera innanzi tutto dalla TM un *fuzzy match S'*, insieme alla sua traduzione *T'*. La TU estratta (*S'* e *T'*) viene quindi utilizzata come metro di confronto per ricercare una traduzione "compatibile" con la TU stessa. Per far ciò viene sfruttata la capacità del sistema di TA di recuperare porzioni di frase non contigue o separate (*bi-phrases*), a partire non solo da *S* ma anche da *S'* e *T'*, e in particolare dalla porzione di frase in comune tra *S* e *S'* e dalla sua corrispondente traduzione recuperata dalla TA. Il sistema di TA utilizzato, statistico e *phrase-based*, è stato prima arricchito con una lista di espressioni formate da parole contigue (es. "catch fire") e non (es. "not ... anymore"), estratte da un corpus parallelo e allineate ai loro corrispondenti nella lingua di arrivo secondo una metrica parola per parola (cf. 5.4). Ad esso è stata poi affiancata una TM già popolata e contenente dei *fuzzy match* utilizzabili nel testo da tradurre. La TM dinamica recupera per prima i principali candidati per la traduzione di una frase *S*, ordinati in base al numero di parole in comune (percentuale di match) e successivamente anche in base alla lunghezza della porzione comune (LCS – *Longest Common Subsequence*), in modo tale da creare la traduzione più corretta a partire da questi due risultati. L'ultimo passaggio consiste nel rinvenire la corretta traduzione per i gap, ovvero le parti di testo non corrispondenti all'interno del match, che verranno desunte a partire dal corpus di testi allineati. Tale metodo sembra offrire notevoli miglioramenti nei valori BLEU e NIST, utilizzati per valutare qualitativamente la performance dei sistemi di traduzione automatica (cf. par. 6.4). La sua caratteristica vincente è la capacità di riordinare in maniera accurata le parole e le espressioni composte all'interno della frase tradotta, grazie alla modalità di allineamento che ricerca gli equivalenti per ogni parola anche nel senso inverso alla lingua di partenza (si pensi per esempio alla diversa posizione dell'aggettivo o alla costruzione dei nomi composti in inglese e in<sup>43</sup> italiano). L'identificazione di componenti più piccole

---

<sup>43</sup> Come è noto, nella maggior parte dei casi gli aggettivi qualificativi in inglese precedono il nome a cui si riferiscono, mentre in italiano generalmente lo seguono, pur essendo possibile in molti casi una maggiore flessibilità concessa da motivi stilistici o per dare maggiore rilievo all'aggettivo stesso ponendolo in posizione anticipata. Allo stesso modo, anche i nomi o le espressioni composte in inglese vengono costruite anticipando gli elementi modificatori, sostantivi, aggettivi, preposizioni o persino verbi, rispetto al sostantivo a cui si riferiscono (es. *address book, washing machine, up-rise etc.*), mentre in italiano avviene esattamente il contrario (Advanced Grammar in Use, Cambridge University Press)

della frase è una strategia ottimale per consentire una migliore riorganizzazione nonché una traduzione maggiormente accurata (Biçici e Dymetman, 2008).

Un altro interessante esperimento è stato condotto dalle Forze Armate statunitensi per la creazione di un sistema di TA per delle lingue che non dispongono di grandi risorse, come quelle dei paesi in cui l'esercito statunitense svolge delle missioni (Winkler *et al.*, 2013). In particolare, emerge da un lato il bisogno di una traduzione altamente settoriale e specifica, e dall'altro la scarsa disponibilità di traduttori in grado di svolgere questa mansione. L'intento dell'*Army Research Laboratory* è quello di creare dei sistemi per la traduzione automatica dei vari ambiti settoriali di loro interesse tramite la raccolta e l'allineamento per mezzo di CAT Tools delle traduzioni già svolte e l'inserimento di tali risorse all'interno di un sistema base di TA. Tale sistema viene poi utilizzato per fornire delle "ipotesi di traduzione" di nuovi testi, che verranno revisionate dai traduttori umani e, una volta corrette, reinserite all'interno del sistema di TA. Si verrà a creare in questo modo un circolo virtuoso dal quale il sistema di TA guadagnerà in ampiezza di risorse, accuratezza e specificità.

L'utilizzo congiunto di CAT Tools e TA si è chiaramente rivelato molto promettente. I sistemi così sviluppati mostrano un ottimo potenziale per l'applicazione di nuove metodologie operative. Sebbene negli studi condotti i sistemi utilizzati fossero stati specificamente creati o adattati agli scopi della sperimentazione, e la loro preparazione sia lunga e complicata, l'applicazione delle medesime procedure e dei medesimi criteri all'interno dei sistemi commerciali non specifici rappresenta sicuramente una delle sfide più interessanti per il futuro e una concreta possibilità di miglioramento della qualità delle traduzioni prodotte.

### 6.3 *La traduzione automatica sul web – sistemi a confronto*

#### **6.3.1. Le origini**

In generale, quando si parla di traduzione automatica il pensiero va ai sistemi resi disponibili in maniera gratuita sul web, e si ritiene erroneamente che tale fenomeno sia nato non più di una decina di anni fa. In realtà, i primi sistemi di TA sul web fecero la loro comparsa dalla metà degli anni '80, quando la rivoluzione delle telecomunicazioni e l'avvento di Internet fecero segnare un picco mai registrato nella richiesta di traduzioni

da parte di singoli individui o enti e aziende pubbliche e private. È in questo periodo che vengono commercializzati anche i primi PC e i primi sistemi di traduzione assistita o automatica per i traduttori freelance. I primi servizi di traduzione automatica sul web non erano ovviamente i grandi sistemi di oggi che offrono traduzioni in moltissime lingue e in maniera istantanea. Si trattava di servizi limitati sia dal punto di vista delle combinazioni linguistiche (prevalentemente le lingue europee) che per la lunghezza dei testi che potevano tradurre; il loro utilizzo aveva un costo e, con una spesa aggiuntiva, era possibile richiedere una revisione umana delle traduzioni ottenute (Hutchins, 2003: 18). Questi sistemi erano prevalentemente pensati per società e aziende, con le quali venivano stipulati accordi e sviluppati modelli *custom-built*. Non era ancora il tempo dei servizi aperti a tutti, che costituiscono una vera e propria rivoluzione nella possibilità di offrire in particolar modo agli utenti comuni (non specialisti) un nuovo modo di guardare alle lingue straniere. Il primo servizio online, Minitel, che utilizzava il sistema sviluppato da Systran, comparve in Francia alla fine degli anni '80. Dopo di esso Comuserve nel 1994, e nel 1997 il celebre Babelfish di Altavista (Giammarresi in Di Sparti, 2006: 116 - 117). Questi servizi online erano inizialmente pensati per tradurre pagine web, posta elettronica o discussioni tramite messaggistica istantanea. Riuscivano a gestire il linguaggio a marcatori e le strutture dei documenti sul web, estrapolando da essi il testo da tradurre. Poco per volta i sistemi consentirono di tradurre porzioni sempre più ampie di testo e di caricare documenti diversi dalle pagine web. Si diffuse anche la consuetudine, per gli enti i gestori di siti web, di ospitare all'interno dei loro portali un collegamento ai sistemi di TA, per attirare gli utenti stranieri e consentire loro di tradurre nella propria lingua il contenuto delle pagine visitate, visualizzandolo all'interno della pagina stessa. Anche i motori di ricerca collegati ai diversi browser sono oggi in grado di identificare la lingua di una pagina e richiedere all'utente se desidera che essa venga tradotta dal servizio di TA che essi integrano (Folaron, 2010b). Altre pagine web, solitamente quelle dei maggiori enti o aziende in grado di stanziare fondi più corposi su questo fronte, sono invece disponibili in più lingue con traduzioni prodotte da traduttori umani. Questa scelta è sicuramente più onerosa in termini di spese per chi gestisce questi siti, ma garantisce al contempo una traduzione qualitativamente migliore e che tenga anche conto di rispettare il layout della pagina, consentendo quindi una visualizzazione più funzionale. L'effetto finale è per l'utente quello di stare visitando una pagina scritta in partenza nella lingua di

traduzione, mentre quando si fa tradurre una pagina da un traduttore automatico, la sensazione che l'utente percepirà è quella di stare visitando una pagina spesso disordinata e poco chiara, oltre che mal tradotta. Sebbene alle volte la pur scarsa qualità della TA consenta comunque di comprendere il contenuto del testo, assolvendo quindi alla sua principale finalità comunicativa, essa non potrà fare a meno di suonare innaturale e inaccurata, a detrimento dell'esperienza complessiva dell'utente finale. Tuttavia, il ricorso ai sistemi automatici offre anche dei vantaggi oltre alla relativa convenienza economica. Ad esempio, in genere i gestori delle pagine e dei portali web offrono la traduzione dei loro contenuti in un numero limitato di lingue, mentre gli attuali sistemi di TA possono tradurre potenzialmente verso un numero sempre crescente di idiomi (al momento Google Translate ne offre 92).

### **6.3.2. Sistemi a confronto**

I sistemi gratuiti pensati per un'utenza poco specializzata e sporadica vengono offerti sul web da diverse società, tra le quali anche le produttrici dei software a pagamento. Uno dei più antichi sistemi di TA gratuita, ancora oggi in funzione, è quello offerto da Systran<sup>44</sup>, di per sé una delle realtà più antiche nel settore della traduzione automatica. Esso è uno dei pochi strumenti che adottano ancora una metodologia *rule-based* e consente di svolgere traduzioni automatiche da e verso 14 lingue. Ha una struttura e un'interfaccia molto semplice e il suo utilizzo è molto elementare ([www.systranet.com](http://www.systranet.com)). Maggiori funzionalità e opzioni sono offerte ad utenti registrati o nelle sue versioni a pagamento.

---

<sup>44</sup> Fondata da Peter Toma nel 1968, ha operato in particolare per la Commissione Europea e il Dipartimento della Difesa degli Stati Uniti. È stata la prima società a produrre e lanciare un modello ibrido di traduzione automatica.

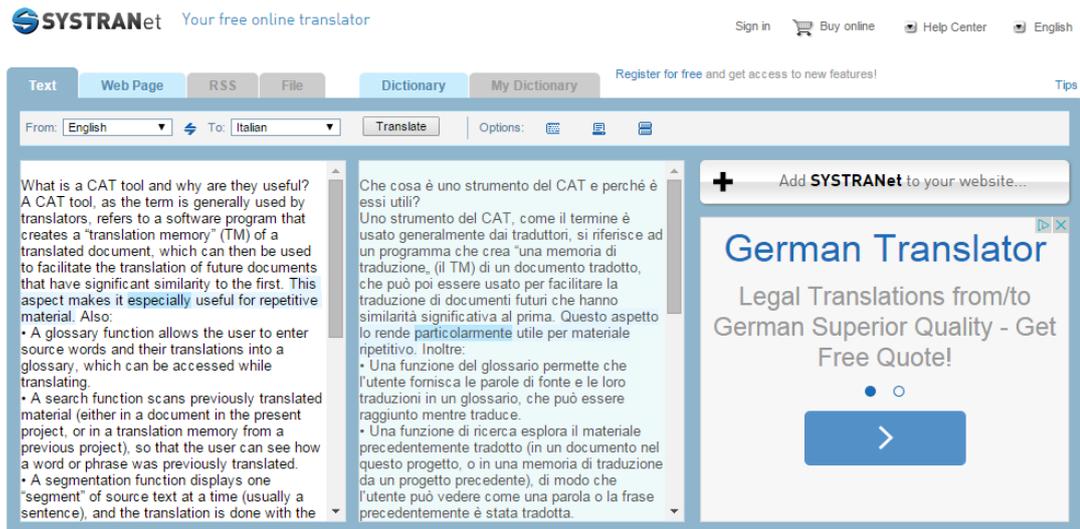


Fig. 27: Il sistema gratuito di traduzione automatica offerto da Systran

Come si vede dall'immagine, Systranet consente di associare frasi e parole del testo di partenza con quelle del testo di arrivo evidenziandole al passaggio del mouse. In questo modo viene facilitata la verifica della traduzione compiuta. Pur provenendo da un passato importante, questo sistema ha subito negli ultimi anni un progressivo declino nel numero dei suoi utenti, a vantaggio dei sistemi statistici che sono quelli maggiormente utilizzati oggi. Si stima che Google Translate abbia mensilmente circa 200 milioni di utenti. Och (2012, cit. in Champsaur, 2013: 21), sviluppatore del sistema statistico adoperato da questo *tool* e attuale capo del dipartimento di Google per la traduzione automatica, stima che ogni giorno venga tradotto tramite di esso l'equivalente di un milione di libri, stima che rende Google Translate il più importante fornitore di traduzioni del mondo. Fino al 2007, come diversi altri sistemi di TA, anche Google si è avvalso di Systran come software di base. Dopo questa data ha adottato invece un software di traduzione proprio di tipo example-based, con caratteristiche innovative.

Google Translate offre numerosi vantaggi, dall'elevato numero di lingue da e verso le quali è possibile tradurre (ad oggi più di 90), alla possibilità di modificare la traduzione ottenuta scegliendo tra più opzioni suggerite dallo strumento stesso o digitando la propria versione corretta e salvandola a beneficio degli utenti futuri:

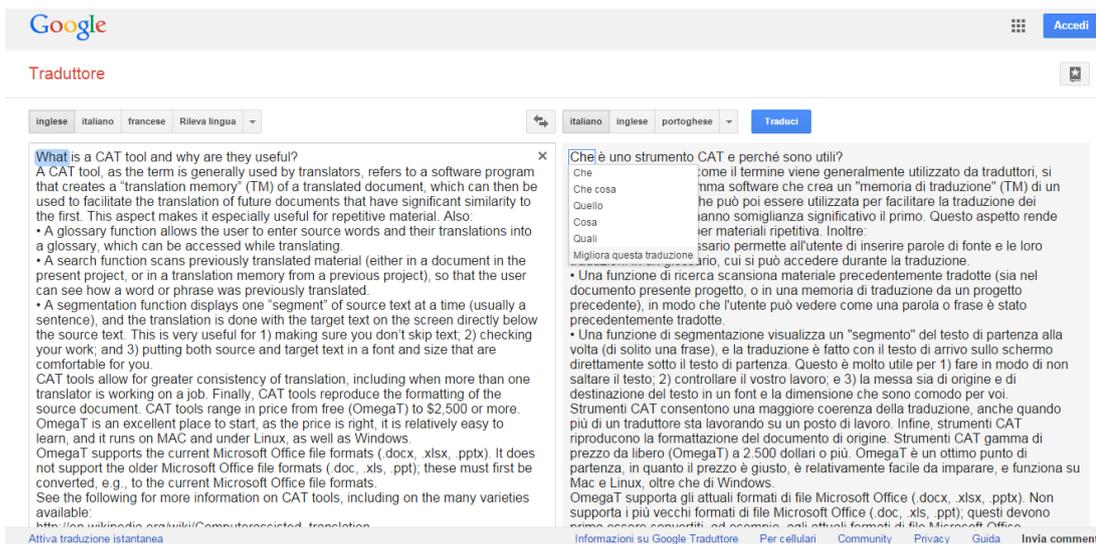


Fig. 28: L'interfaccia di Google Translate e l'opzione di modifica del testo tradotto

La visualizzazione risulta più comoda che nella ristretta interfaccia di Systranet, e la possibilità di migliorare personalmente la traduzione ottenuta rappresenta indubbiamente il valore aggiunto di questo strumento e la chiave del suo continuo accrescimento qualitativo (Thicke, 2013: 16). Trattandosi di un sistema statistico, le traduzioni risulteranno migliori per quelle lingue che possono contare su di un'ampia raccolta di risorse di riferimento. Le lingue più rare e introdotte di recente non garantiscono probabilmente una grande correttezza di forma e contenuti, tuttavia il sistema ha il merito di includerle e migliorare progressivamente la loro implementazione con i contributi dei fruitori. Oltre alla correzione delle traduzioni ottenute è infatti possibile per gli utenti collaborare in maniera più strutturata al miglioramento del servizio tramite la *Translate Community*. Gli utenti registrati potranno tradurre frasi e parole verso la loro lingua, valutare le traduzioni prodotte da Google Translate o inserire alternative per la formulazione di frasi ed espressioni<sup>45</sup>. Se da un lato questa forma molto democratica di partecipazione può porre il dubbio dell'affidabilità dei contributi forniti, essa rappresenta tuttavia una interessante modalità *user-based* di valutazione e miglioramento della performance di tale sistema.

Tra gli altri sistemi di TA più utilizzati del web va anche citato Bing Translator, un sistema statistico sviluppato da Microsoft che offre traduzioni da e verso 45 lingue e alcune delle funzioni che troviamo anche in Google Translate. Per esempio, inserendo

<sup>45</sup> [www.translate.google.com/community?source= localized](http://www.translate.google.com/community?source=localized)

nell'interfaccia di traduzione non un testo ma una singola parola, oltre alla traduzione verrà visualizzata anche la voce di dizionario con delle informazioni ed esempi di utilizzo della parola tradotta:

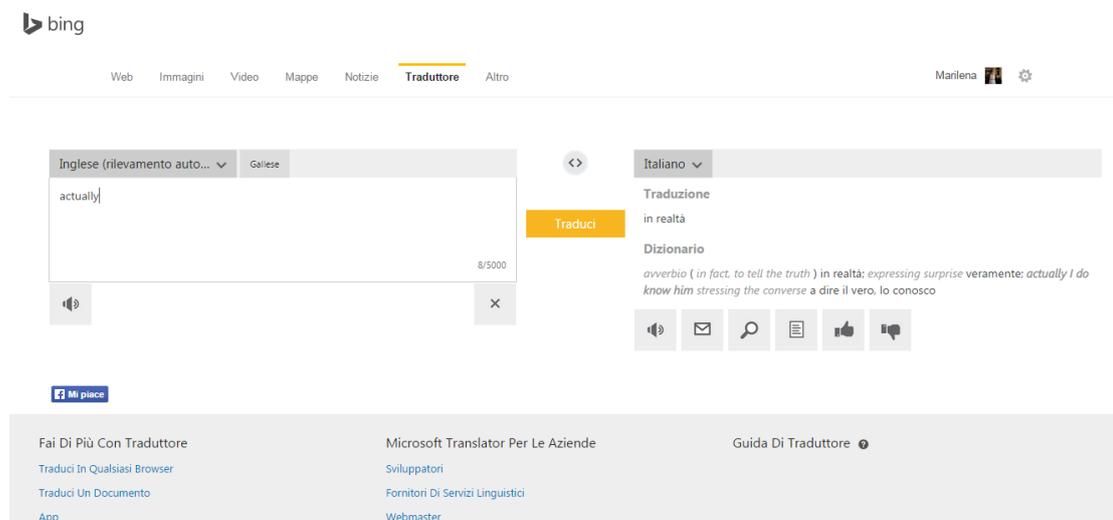


Fig. 29: L'interfaccia di Bing Translator

Come anche gli altri strumenti descritti, Bing Translator ha una capacità limitata in termini di lunghezza di testo traducibile (nel caso specifico 5000 caratteri per volta), e sempre come gli altri strumenti, offre la possibilità di invertire la direzione di traduzione, visualizzare la corrispondenza tra parole e frasi del testo di origine e di quello tradotto, e infine può rilevare automaticamente la lingua del testo di partenza (questa opzione non è supportata da Systranet). Una differenza non immediatamente visibile all'occhio ma che può rivelarsi importante nella fase successiva alla traduzione è rappresentata dal tipo di spaziatura inserita tra le parole. Solo Google Translate inserisce nel testo tradotto un vero e proprio spazio per separare una parola da un'altra e l'interruzione di frase in corrispondenza di un *a capo*. Sia Systranet sia Bing Translator inseriscono invece degli spazi unificatori, i "cerchietti" che vengono rivelati dallo strumento mostra/nascondi quando si copia e incolla testo dal web, e al posto dell'interruzione di frase inseriscono un ritorno a capo, che non divide le due frasi ma le lega insieme. Questo può essere un problema nel caso in cui si decida di revisionare il testo utilizzando un CAT Tool. Le frasi non divise andranno a formare un unico blocco che non permetterà la visualizzazione allineata al testo di partenza. Anche nel caso in cui il testo venga direttamente inserito sul foglio di Word si presenteranno problemi a causa degli spazi mancanti, con righe

interrotte e difficoltà di rispettare l'impostazione grafica della pagina. In ultima istanza, occorrerà sostituire tutti gli spazi unificatori con degli spazi veri e propri, operazione frustrante e che comporterà un ingente investimento di tempo in base alla lunghezza del documento.

### 6.4.3. Traduzioni a confronto

Venendo invece al tipo di traduzione che si può ottenere con i sistemi sopra elencati, si prenderà ad esempio un piccolo stralcio di testo tradotto dall'inglese all'italiano<sup>46</sup> dai tre sistemi appena descritti. Occorre riconoscere che in tutti e tre i casi la traduzione prodotta consente di comprendere senza troppe difficoltà il contenuto del testo, ma laddove si dovesse utilizzare la traduzione ottenuta per un qualsiasi scopo che ne preveda la condivisione, pubblicazione ecc., in nessuno dei tre casi essa potrebbe essere adottata così come prodotta dagli strumenti, ma necessiterà comunque di una revisione accurata da parte dell'utente.

Originale	Systranet	Bing Translator	Google Translate
What is a CAT tool and why are they useful?	Che cosa è uno strumento del CAT e perché è essi utili?	Che cosa è uno strumento CAT, e perché sono utili?	Che è uno strumento CAT e perché sono utili?

Tab. 1: La traduzione del titolo del documento

Questo primo esempio è già abbastanza significativo della qualità delle traduzioni ottenute. Systranet, nel suo tentativo di trasporre le regole grammaticali e sintattiche della lingua di partenza nelle corrispondenti regole della lingua di arrivo, inserisce la preposizione articolata *del* con valore possessivo, e ancor più bizzarramente accosta un verbo alla terza persona singolare con un soggetto plurale (*è essi*), creando una traduzione comprensibile ma che suona innaturale a chi la legge. Le traduzioni di Bing e Google non presentano veri e propri errori, se non per l'eccepibile "Che è" prodotto da Google Translate. Malgrado ciò, esse non sono comunque corrispondenti alle aspettative che un traduttore o un utente specializzato potrebbe avere. Intanto il termine "CAT Tool" viene utilizzato anche in italiano, preferibilmente a "strumento CAT" che comunque non è

---

<sup>46</sup> L'originale è stato estratto dal manuale di utilizzo di OmegaT, il più famoso CAT Tool open source al quale si è accennato in 2.3.

sbagliato. Quello che suona strano in italiano è l'accostamento di una frase al plurale successivamente ad una con soggetto al singolare. Un traduttore avrebbe probabilmente trasformato la prima parte adoperando un soggetto plurale per creare maggiore accordo con la seconda parte del titolo (Che cosa sono i CAT Tool e perché sono utili?). Le cose si complicano con una porzione di testo leggermente più ampia:

Originale	Systranet	Bing Translator	Google Translate
A CAT tool, as the term is generally used by translators, refers to a software program that creates a "translation memory" (TM) of a translated document, which can then be used to facilitate the translation of future documents that have significant similarity to the first. This aspect makes it especially useful for repetitive material.	Uno strumento del CAT, come il termine è usato generalmente dai traduttori, si riferisce ad un programma che crea "una memoria di traduzione,, (il TM) di un documento tradotto, che può poi <u>essere usato</u> per facilitare la traduzione di documenti futuri che hanno similarità significativa al prima. Questo aspetto lo rende particolarmente utile per materiale ripetitivo.	Uno strumento CAT, come il termine è generalmente usato dai traduttori, che si riferisce a un programma che crea una "memoria di traduzione" (TM) di un documento tradotto, che quindi può <u>essere utilizzato</u> per facilitare la traduzione di documenti futuri che hanno somiglianza significativa al primo. Questo aspetto lo rende particolarmente utile per materiale ripetitivo.	Uno strumento CAT, come il termine viene generalmente utilizzato da traduttori, si riferisce ad un programma software che crea un "memoria di traduzione" (TM) di un documento tradotto, che può poi <u>essere utilizzata</u> per facilitare la traduzione dei documenti futuri che hanno <u>somiglianza significativa</u> il primo. Questo aspetto rende particolarmente utile per <u>materiali ripetitiva</u> .

Tab. 2: Un altro stralcio di traduzione

Appare chiaro anche dalla traduzione di questo periodo che la comprensione del contenuto è comunque garantita da tutti e tre gli strumenti. Ognuno di essi produce un testo contenente principalmente degli errori di accordo tra articolo e nome o nome e aggettivo, anche il sistema *rule-based* (Systranet) che dovrebbe in teoria garantire maggiore accuratezza da questo punto di vista; in realtà, la complessità pur ridotta del testo originale non consente una corretta analisi della struttura sintattica e grammaticale, motivo per il quale i sistemi RB sono spesso utilizzati con linguaggi controllati la cui

sintassi lineare limita il numero di fraintendimenti. In particolare è interessante vedere come Google Translate sia l'unico a comprendere che a poter "essere utilizzata" è la memoria di traduzione, mentre Systranet e Bing Translator declinano "usato" o "utilizzato" al maschile perché lo associano a "documento". Tuttavia, l'intuito di Google viene penalizzato da quei "somiglianza significativo" e "materiali ripetitiva" che denotano una difficoltà dei sistemi statistici di risolvere anche i più piccoli problemi di accordo degli aggettivi.

Non si vuole qui determinare quale degli strumenti descritti sia il migliore, perché come si è visto nessuno di essi spicca rispetto agli altri per una evidente maggiore qualità del testo ottenuto. In genere, le traduzioni prodotte dai sistemi RB risultano più corrette dal punto di vista grammaticale, mentre quelle ottenute dai sistemi statistici risultano più scorrevoli e comprensibili (Hutchins, 2003), tuttavia come emerge chiaramente dal confronto svolto, nessun sistema è in grado di ridurre significativamente la presenza di errori o imprecisioni. Piuttosto che sforzarsi di valutare quale sistema fornisca la traduzione "meno cattiva" occorre principalmente comprendere che i sistemi di traduzione automatica gratuita offerta sul web hanno dei limiti, e vanno utilizzati in maniera critica. Essi possono aiutare sicuramente nella comprensione di documenti scritti in lingue straniere, e in tal senso sono un ottimo strumento (Hutchins, 2003: 23). Tuttavia, a prescindere da qualsiasi strumento si adotterà, non ci si potrà comunque mai aspettare una traduzione formalmente corretta, priva di errori e dalla qualità sufficientemente buona da essere utilizzata così come è prodotta dal computer. Laddove la TA venga utilizzata per finalità diverse dalla semplice comprensione di un testo, essa dovrà essere sottoposta ad un'attenta revisione, o si dovrà optare per strumenti e servizi diversi in grado di garantire il rispetto gli standard perseguiti.

## *6.4 Valutazione dei sistemi di traduzione automatica*

### **6.4.1. Punti di vista diversi**

Per concludere questa trattazione occorre soffermarsi su un aspetto che può rivelarsi molto interessante per comprendere quanto varia e differenziata sia la branca di studio relativa alla traduzione automatica e quanti fattori e punti di vista diversi convergano in essa. Si tratta dei metodi e degli strumenti sviluppati dai ricercatori al fine

di valutare la performance e l'evoluzione dei sistemi di TA. La misurazione o valutazione è un aspetto centrale in ogni ambiente scientifico. È attraverso delle misurazioni accurate che è possibile apprezzare in maniera chiara e concreta i progressi ottenuti, ed è necessario che tali valutazioni siano anche predittive per poter avanzare ipotesi sensate e concrete sui potenziali risvolti futuri a partire dai risultati reali ottenuti (Graham et al. 2012). Tali metodi, come è facile intuire, variano in base alle diverse prospettive adottate dagli studiosi a seconda del proprio campo di interesse: gli sviluppatori dei software si concentreranno maggiormente sull'usabilità dei loro sistemi, i linguisti sulla qualità della traduzione prodotta, gli studiosi delle intelligenze artificiali sulla capacità dei sistemi di riprodurre i meccanismi del linguaggio umano. È evidente che le prospettive e gli obiettivi che i diversi specialisti hanno in mente sono per forza di cose diversi, ma che ognuno di essi ha la sua importanza e insieme consentono di comprendere e giudicare il fenomeno nella sua interezza. È quindi necessario, nel valutare, tenere conto di tutti i diversi livelli e aspetti dei *machine translation tools*. Storicamente le prime valutazioni furono di tipo linguistico, esclusivamente legate al confronto tra la traduzione prodotta dal computer e la traduzione umana, come nel caso di quelle condotte per la compilazione del rapporto ALPAC del 1966 a cui si è già fatto riferimento. Tale valutazione era tuttavia pressoché insignificante perché partiva dal presupposto di poter uguagliare due processi in realtà inaccostabili sotto molti punti di vista. Inoltre, che una traduzione automatica non potesse essere buona alla pari di una traduzione umana è una conclusione decisamente scontata, e il fatto che il rapporto ALPAC la sottolinei con enfasi fa comprendere quali grandi aspettative avessero accompagnato il lancio dei primi sistemi di TA negli anni Cinquanta. Successivamente gli studiosi iniziarono ad interrogarsi maggiormente sui criteri valutativi, nell'ottica dell'ottenimento di risultati utili al miglioramento dei sistemi sviluppati, i quali iniziarono ad essere misurati in termini di qualità del prodotto e di utilità all'interno di reali contesti operativi (contestualmente al periodo in cui sorsero gli studi sui linguaggi controllati con il riconoscimento della migliore qualità dei risultati per contesti linguistici chiaramente determinati e ristretti). Infine nacquero i primi studi comparativi tra sistemi, svolti sia da centri di ricerca che da potenziali acquirenti (Hutchins, 1997), anche in questo caso con finalità diverse ma ugualmente utili a fornire un panorama esaustivo sul fenomeno. Indubbiamente, a contribuire alla formulazione delle metodologie di valutazione furono gli studi sul NLP (il trattamento dei linguaggi

naturali o *Natural Language Processing*), all'interno dei quali vengono distinte tre tipologie di analisi. La prima è relativa all'adeguatezza dei sistemi a specifici contesti operativi, la seconda alla presenza di limiti, errori e imperfezioni che possono essere corretti o migliorati dal gruppo di ricerca o dagli sviluppatori del software in base alle diverse competenze, e infine la terza è relativa alla performance dei sistemi implementati (Hutchins, 1997: 418). Queste tre iniziali articolazioni si applicano anche ai sistemi di TA, rappresentando i criteri generali per la loro valutazione. A partire da questi sono stati identificati altri parametri peculiari per la valutazione specifica dei sistemi di TA, ovvero la qualità della traduzione "grezza" (non revisionata), l'usabilità delle varie funzioni offerte dai sistemi, l'estendibilità a nuove combinazioni linguistiche e o domini specifici e infine il confronto tra costi e benefici rispetto alla traduzione umana (*ibid.*).

#### **6.4.2. Metodologie di valutazione umana**

Come è evidente, la gamma di aspetti e di variabili da valutare è ampia oltre che molto varia, e necessita quindi di strumenti specifici per i vari parametri. Una prima distinzione delle diverse metodologie di valutazione esistenti per la TA è quella tra i criteri basati sul giudizio umano e quelli basati su parametri che possono essere raccolti e valutati in maniera computerizzata (Graham et al. 2012). I primi prevedono la raccolta di giudizi espressi da diversi utenti, preferibilmente specializzati nell'ambito in esame, e il loro confronto per stabilire un giudizio medio condiviso. I parametri che è possibile valutare in questo modo sono molteplici, ma in particolare si ricorre al giudizio umano quando si vuole ottenere una valutazione qualitativa della traduzione prodotta o svolgere un'analisi comparativa tra più sistemi. Il giudizio qualitativo prevede di considerare alcuni aspetti che la macchina non potrebbe cogliere, quali l'adeguatezza (la chiarezza e correttezza nell'espressione dei concetti) o la scorrevolezza (facilità di lettura data dalla forma più o meno corretta) del testo (Kohen, 2004), o prendere in considerazione aspetti specifici come le unità lessicali, la struttura sintattica ecc. (Papineni *et al.*, 2002). Tra le metodologie più comuni relative alla valutazione umana delle traduzioni automatiche si possono annoverare l'osservazione frase per frase con la possibilità di inserire commenti o dare valutazioni su una scala numerica, il confronto tra diverse traduzioni finalizzato a stabilire una scala gerarchica in termini di qualità e rispondenza alle aspettative, nonché il tempo impiegato per il post-editing delle traduzioni prodotte dai sistemi di TA (Doherty, O'Brien, Carl, 2010). I giudizi umani sono anche utilizzati come misura per

valutare l'attendibilità e affidabilità delle misurazioni automatizzate, sviluppate all'inizio come possibili alternative alle lunghe e costose valutazioni umane.

### **6.4.3. Metodologie di valutazione automatizzata**

Il primo obiettivo dei sistemi automatizzati di valutazione era quello di determinare il grado di somiglianza tra traduzioni prodotte dal computer e una traduzione umana di riferimento (Graham et al. 2012), assegnando un punteggio a ciascuna e ordinando le traduzioni automatiche sulla base dei punteggi ottenuti. Tuttavia questo metro di valutazione si rivelò presto poco utile, sia perché anche gli equivalenti traduttivi, pur se corretti, venivano computati come erronei semplicemente perché diversi dal testo di riferimento, sia per la difficoltà di valutare la somiglianza sintattica tra due traduzioni, e soprattutto di determinare quale, tra una serie di traduzioni automatiche, fosse la più simile alla traduzione di riferimento. Nel complesso, questo sistema non forniva alcuna informazione sulla reale qualità della traduzione, e predisponeva al rischio di un lavoro di perfezionamento dei software finalizzato più ad ottenere un punteggio elevato che ad un reale miglioramento della qualità delle traduzioni. Tuttavia, la necessità di ottenere rapide indicazioni sulla performance dei sistemi ha spinto i ricercatori a puntare sui metodi di valutazione automatizzata, per sviluppare dei sistemi semplici, poco costosi e applicabili a prescindere dalle combinazioni linguistiche coinvolte.

Uno dei metodi più diffusi, a cui si è già fatto riferimento in 6.2, è il punteggio BLEU (*bilingual evaluation understudy*). Esso parte sempre dall'idea di poter valutare una traduzione prodotta automaticamente sulla base del suo grado di somiglianza con una traduzione umana di riferimento, ma tale criterio viene ampliato introducendo più traduzioni di riferimento e un algoritmo di estrazione degli *n*-grammi (porzioni di testo) comuni tra le varie traduzioni da valutare e quelle di riferimento, capace di prendere in considerazione diverse traduzioni per volta e ordinare le loro performance secondo una scala qualitativa (Papineni *et al.*, 2002; Kohen, ). Il sistema BLEU è uno dei più diffusi negli studi sui nuovi modelli di TA, per la semplicità di utilizzo e la rapidità con la quale esso fornisce i dati valutativi. Alcune importanti obiezioni possono tuttavia essere mosse a tale sistema. Prima di tutto, stabilire quale sia la migliore traduzione automatica tra un campione preso in esame non implica direttamente che essa sia una traduzione di buona qualità. Un elevato tasso di somiglianza tra due frasi non è sempre e comunque sinonimo di una loro corrispondenza a livello tanto di forma quanto più significativamente di

contenuto, quindi, per quanto una traduzione possa essere valutata molto prossima a un set di traduzioni di riferimento, ciò non potrà fornire una certezza di qualità se non in maniera congiunta con una valutazione umana. Indubbiamente il sistema BLEU può fornire delle utili indicazioni sulla performance di sistemi diversi, e le sue valutazioni risultano corrette e coerenti se confrontate con delle valutazioni umane (*ibid.*). Il metodo adottato e l'obiettivo perseguito lasciano comunque pensare a tale sistema come ad un metro valutativo più adatto agli sviluppatori di software piuttosto che agli studiosi della traduzione automatica come fenomeno linguistico.

Una maniera forse più prossima al punto di vista dei linguisti e dei traduttori è quella già accennata che analizza la qualità delle traduzioni automatiche sulla base dei tempi richiesti per la revisione (o post-editing). Laddove infatti la TA venga impiegata dai traduttori professionisti, sarà per uno scopo che richiede in ogni caso uno standard qualitativo molto alto, quindi la revisione e l'adeguamento del testo saranno sempre necessari. Una traduzione "grezza" di qualità accettabile avrà tempi di revisione più brevi rispetto a una traduzione di cattiva qualità (Doherty, O'Brien, Carl, 2010). Questo metro di valutazione offre inoltre il beneficio di avvalersi di figure specializzate – traduttori o revisori – capaci di fornire giudizi affidabili data l'esperienza professionale maturata. Al fine di determinare la complessità del processo di post-editing in alcuni studi si è fatto ricorso in via sperimentale a una metodologia già incontrata nel corso di questo lavoro (cf. 3.5), ovvero l'*eye tracking*, il monitoraggio dell'attività oculare, significativo secondo l'ipotesi dell'interrelazione occhio-mente (*ibid.*). Secondo tale ipotesi, quando l'occhio si sofferma su di un oggetto, per esempio una frase, il cervello è impegnato in qualche tipo di elaborazione cognitiva del dato oggetto. Ad ogni pausa nei movimenti oculari (fissazione) corrisponde in genere una criticità che il cervello deve elaborare al fine di determinarne la soluzione (O'Brien, 2006). Numerose fissazioni implicano dei tempi totali di revisione più lunghi, che possono essere rivelatori di una traduzione automatica di cattiva qualità. Tuttavia, al di là della qualità della traduzione, della cui inevitabile imperfezione ogni traduttore o revisore sarà ben consapevole senza il bisogno di alcuna valutazione, il fattore tempo è di per sé determinante nel mercato della traduzione. Alcuni studi affermano che revisionare una traduzione prodotta automaticamente comporti un risparmio del 50% del tempo che sarebbe impiegato per tradurre il testo ex novo (Mossop, 2007); un sistema di TA capace di fornire delle traduzioni "grezze" revisionabili in tempi

brevi – quindi non eccessivamente piene di errori – sarà la scelta preferenziale per agenzie e traduttori. La prospettiva che soggiace a questo metodo, riflettente il punto di vista e le aspettative di questa categoria di professionisti, è chiaramente diversa da quella soggiacente allo sviluppo delle metodologie automatizzate. Più che valutare quanto la traduzione si avvicini alla qualità della traduzione umana, essa intende misurare se e quanto lo strumento traduzione automatica possa essere materialmente utile ai traduttori e in che misura possa semplificarne o accelerarne il lavoro.

È chiaro, quindi, che occorre distinguere e comprendere le diverse metodologie di valutazione e gli obiettivi che ognuna di essa persegue. La molteplicità di approcci e prospettive che caratterizzano il fenomeno della traduzione automatica trova conferma ed è messa in particolare luce dalla questione valutativa, sia nella sua dimensione positiva di interdisciplinarietà, che in quella più critica di incommensurabilità dei punti di vista e delle finalità ultime. Sebbene i giudizi umani e le valutazioni automatizzate possano trovare dei punti di contatto, e integrarsi nel tentativo di fornire una misurazione esaustiva e completa degli strumenti di TA, è utile guardare ad essi in maniera analitica e saper distinguere i diversi livelli coinvolti per una migliore comprensione del fenomeno.

## Capitolo 7: Una riflessione critica

### *7.1 Il punto di vista dei traduttori*

L'ultimo capitolo di questa Seconda Parte tenterà di mettere insieme i vari “pezzi” del poliedrico fenomeno della traduzione automatica, per compiere una riflessione critica che addotti il duplice punto di vista di chi scrive, ovvero quello proprio dei traduttori e degli studiosi di traduzione afferenti al ramo “applicato” dei Translation Studies secondo il già citato schema di Holmes (cf. 1.5). La prima di queste due categorie di esperti è quella che, insieme ai teorici “classici” della traduzione, ha storicamente avuto un rapporto conflittuale con la traduzione automatica, fondamentalmente per il timore delle ricadute negative di tale strumento sul lavoro e sullo status sociale della figura professionale del traduttore (Hutchins, 2003). La seconda categoria racchiude invece quel gruppo di studiosi che hanno guardato e guardano ancora oggi alla traduzione automatica con interesse in quanto fenomeno sociale oltre che linguistico, ma con il giusto grado di prudenza che manca delle volte agli esperti degli ambiti più tecnici. È stato proprio questo il punto di vista che si è tentato di adottare nel corso di questa trattazione, nella convinzione che non si possa liquidare la traduzione automatica con una superficiale considerazione sulla sua inadeguatezza e dannosità, ma che se ne debbano riconoscere meriti e demeriti, ed ancor più le potenzialità di sviluppo e i margini di un suo utilizzo reale, data la sua ormai indiscussa affermazione sul mercato e nella vita di ogni giorno.

#### **7.1.1. Il questionario d'indagine**

Al contributo e alla posizione dei traduttori si è solo fatto qualche breve accenno nel corso dei capitoli, nel tentativo di identificare i motivi della loro generale opposizione al dilagante uso dei sistemi di TA. Si è inoltre evidenziato il ruolo marginale che è stato concesso a linguisti e traduttori nella fase di sviluppo dei software (Pym, 2006), e la poca voce in capitolo che essi hanno sempre avuto nella programmazione ed applicazione dei

sistemi di TA, come testimoniato dalle parole di Beesley (1986, cit. in Taravella e Villeneuve, 2013: 64):

*The most neglected aspect of MAT [machine-assisted translation] is perhaps the most important: the needs, attitudes and sensitivities of human translators. MAT researchers often ignore the fact that human beings have to use their systems [...] The cooperation of human translators is [...] essential to make any MAT installation work. If translators are saddled with intimidating and inappropriate technology or if the machines become their masters rather than their slaves, then the technology will fail – the translator will see to it.*

Il ruolo marginale sempre riservato ai traduttori nella ricerca sulla TA è il principale motivo per cui essi hanno sviluppato un atteggiamento di opposizione netta nei confronti di tale strumento, considerato da molti di loro un surrogato di competenza linguistica, un rivale, un pretesto delle agenzie di traduzione per sottopagare la professionalità dei traduttori. Queste considerazioni trovano conferma nelle dichiarazioni raccolte tramite un questionario appositamente preparato da chi scrive per cogliere alcuni significativi aspetti del rapporto che i traduttori hanno nei confronti degli strumenti digitali per la traduzione e in particolar modo della traduzione automatica. Il questionario è stato sottoposto ad un gruppo di 30 traduttori professionisti con esperienza nel settore di minimo tre anni. Si è strutturato in una serie di domande a risposta aperta di natura ora più generica ora più specifica, relative all'esperienza generale degli intervistati (anni di lavoro, principali settori per i quali si traduce etc.), al loro rapporto con le tecnologie dedicate alla traduzione e ad alcune considerazioni di principio<sup>47</sup>. Per esempio, è stato chiesto ai traduttori se utilizzassero un CAT Tool in maniera più o meno costante e se facessero ricorso a dei servizi di traduzione automatica, con che frequenza e per quale utilità. È stato anche chiesto di esprimere il proprio parere generale sugli strumenti di TA a partire dall'esperienza personale maturata, e sul modo in cui le tecnologie influenzano il lavoro del traduttore.

---

<sup>47</sup> L'intero questionario è riportato in appendice.

### 7.1.2. Il rifiuto della TA

Relativamente all'utilizzo dei CAT Tools, la quasi totalità degli intervistati ha detto di farne un uso regolare (27 su 30), in alcuni casi di conoscerne ed utilizzarne più tipologie e di farlo per comodità personale (18 su 30) oltre che su richiesta delle agenzie. Le risposte relative all'adozione dei sistemi di traduzione automatica sono ben diverse. Soltanto 5 dei 30 intervistati hanno ammesso di ricorrere saltuariamente ai sistemi di TA o alla funzione di traduzione automatica disponibile all'interno dei CAT Tools, mentre tutti gli altri hanno risposto di non farne mai uso perché ritengono tale strumento più dannoso che benefico<sup>48</sup>. Alcuni dicono di utilizzare come unica forma di traduzione automatica quella basata su una TM, ovvero l'inserimento automatico degli *exact match* reperiti dalle memorie di traduzione create personalmente (riferito da 3 degli intervistati). La principale motivazione fornita alla quasi unanimità per il rifiuto dei sistemi automatici è la scarsa qualità dei risultati ottenuti, che necessitano “non di una revisione, ma di una vera e propria ristesura” (int. N° 12), non presentando quindi alcun vantaggio né dal punto di vista del tempo impiegato che da quello del carico di lavoro sostenuto. Circa un terzo degli intervistati (che lavorano come editor oltre che come traduttori) ha fatto riferimento all'abitudine di alcune agenzie di traduzione di offrire lavori da revisore (la cui retribuzione è pari a circa la metà di quella dei traduttori) per testi tradotti automaticamente la cui qualità iniziale è spesso molto bassa e richiedente circa il doppio del lavoro rispetto alla normale revisione di un testo tradotto da un collega. Un intervistato scrive “questo tipo di richiesta è molto frequente, se il testo è troppo cattivo mi rifiuto perché economicamente non ne vale la pena (...) alle agenzie con cui lavoro più stabilmente richiedo sempre una tariffa maggiore per questo tipo di revisione, il carico di lavoro è quasi pari a dover tradurre tutto da capo” (int. n° 18). Diversi riconoscono che il ricorso alla TA è di unica convenienza delle agenzie di traduzione, e lamentano questa abitudine come offesa alla professionalità dei traduttori e pretesto per una inferiore retribuzione. Ciò trova riscontro in Thicke (2013) che riconosce come la cattiva qualità della traduzione automatica causi un aggravio di lavoro al post-editor, innescando un sistema non sostenibile, sia per la discrepanza tra la remunerazione del revisore ed il suo carico di lavoro, sia per i tempi che richiede e che rendono in ultima istanza il ricorso alla

---

<sup>48</sup> “La pre-traduzione basata su un sistema automatico genera spesso confusione e porta a commettere errori, per esempio a dimenticare di tradurre aggettivi o avverbi, soprattutto quando si ha fretta” (int. N° 3).

traduzione automatica non conveniente. Per questo motivo i traduttori disposti a lavorare come revisori sono sempre di meno, a fronte della crescente richiesta causata dall'applicazione massiccia della TA.

Fondamentalmente, le due motivazioni dietro al rifiuto della traduzione automatica da parte dei traduttori sono quindi di natura pratica ed economica. In un paio di questionari, tuttavia, si è potuto cogliere anche un riferimento ad un'altra motivazione, rappresentata dal danno "d'immagine" che tale strumento può arrecare alla figura professionale del traduttore. Pur trattandosi di un numero esiguo di intervistati, è interessante notare come alcuni traduttori si sentano danneggiati dal dilagante ricorso alla traduzione automatica sia per la diffusione di cattive traduzioni prodotte automaticamente specialmente sul web – sinonimo di una prevalenza dell'aspetto economico su quello qualitativo per molti utenti – sia per la convinzione dell'utenza comune di poter equiparare traduzione umana e traduzione automatica, e di potere fare a meno della prima avendo a disposizione degli strumenti gratuiti e veloci. Afferma un intervistato "Molti non comprendono il carico di lavoro che c'è dietro ad una traduzione. Ho avuto spesso a che fare con clienti privati convinti di poter ottenere traduzioni molto lunghe e complesse in pochi giorni e per costi irrisori. Quando presento il mio preventivo capita spesso che mi si risponda "ma se la faccio con Google non è lo stesso?" (int. n° 9); o ancora " (...) la tecnologia, se utilizzata bene, è molto utile, e in linea di principio non ho nulla contro l'uso della traduzione automatica in certe situazioni. Il problema è fare capire ai clienti l'abissale differenza che separa la traduzione di un professionista da quella automatica" (int. n° 24). A parziale confutazione di questi punti di vista un altro intervistato più ottimisticamente afferma che "la traduzione automatica non è uno strumento per traduttori, né per clienti che vogliono ottenere una traduzione di qualità e intendono avviare un lavoro continuativo con le agenzie. Non penso che noi traduttori possiamo essere danneggiati dalla traduzione automatica, a venirne danneggiati sono caso mai quelli che continuano a farne uso" (int. n° 16).

### **7.1.3. La percezione delle tecnologie**

Relativamente all'impatto delle tecnologie sul loro lavoro, si è rivelata opinione comune degli intervistati la loro utilità nella gestione ed organizzazione delle risorse, nel più agevole reperimento delle risorse utili, nella velocizzazione di alcuni compiti e nella condivisione del lavoro e delle risorse. Le tecnologie più utilizzate sono, oltre ai CAT

Tools, i glossari e internet, utilizzato “per informarmi, più raramente per cercare traduenti” (int. n° 3); “per effettuare ricerche e riscontri terminologici” (int. n° 11); per consultare “dizionari online, siti di traduttori, motori di ricerca” (int. n° 20). Come già detto, la traduzione automatica non viene utilizzata dalla stragrande maggioranza degli intervistati, ma viene lamentato l’impatto negativo dovuto alla sua adozione da parte di alcune agenzie ed alcuni clienti. Alcuni intervistati riconoscono quale impatto negativo della tecnologia la continua produzione di nuove versioni dei software e nuovi strumenti e la difficoltà di mantenersi al passo con gli ultimi lanci sul mercato, nonché l’investimento economico che essi richiedono. Interessante, per concludere, il commento di un intervistato che dichiara “le tecnologie sono necessarie ma alle volte vengono sovraccaricate d’importanza. Un traduttore deve innanzi tutto essere capace di tradurre, le altre conoscenze sono funzionali a questo scopo. Le tecnologie sono solo strumenti e penso debbano essere utilizzate in base all’utilità che ognuno riconosce di trarne” (int. n° 28).

#### **7.1.4. La TA strumento per i traduttori?**

Dai risultati del questionario appare chiaro che il rapporto dei traduttori con i sistemi automatici sia lontano dagli entusiasmi dei ricercatori e degli sviluppatori dei software. La traduzione automatica è per la maggior parte dei partecipanti al questionario una realtà del mercato che occorre conoscere ma non necessariamente utilizzare, per la qualità mediocre dei suoi risultati che necessitano di profonde e lunghe revisioni. Vengono preferite a lei quelle tecnologie che offrono una maggiore sicurezza dei risultati prodotti ed un reale contributo al lavoro. Sebbene le aziende produttrici di software CAT si siano attrezzate a produrre anche dei sistemi personalizzabili di traduzione automatica, e malgrado i progressi significativi ottenuti nel corso del tempo, i traduttori rimangono volutamente estranei a queste dinamiche, a difesa della qualità inarrivabile delle traduzioni umane e di una professionalità che conta su strumenti che ne possano accrescere gli standard e la performance. Come però suggerisce Champsaur (2013), i traduttori non possono ignorare l’esistenza di tale strumento, né sottovalutare la sua continua evoluzione e i miglioramenti raggiunti. La traduzione automatica può diventare uno strumento per i traduttori se utilizzata in maniera critica, in complementarità con altri strumenti come le TM e nella consapevolezza dei suoi limiti. Un’ultima considerazione viene presa in prestito da Melby (1981), relativamente alla questione della

presunta influenza negativa della TA sulla figura professionale del traduttore. Secondo l'autore, la diffusione dei numerosi sistemi di sussidio digitale alla traduzione avrebbe in ultima istanza accresciuto lo status della professione di traduttore, piuttosto che danneggiarlo come temuto da molti<sup>49</sup>. Anche più di recente (2006), come lanciando una sfida ai traduttori, Melby invita ancora a valutare in termini positivi il possibile contributo della traduzione automatica nel processo qualitativo di produzione di:

*This is actually good news for translators: they will be more human, rather than less. They will be involved in the entire quality assurance (QA) process of creating specifications appropriate for the audience and purpose of a particular translation and making sure they are adhered to at every step of the project. If that sounds like a project manager, so be it. The future is ours: do we want to be viewed more as file clerks afraid of being replaced by document management systems or as confident professionals who are gradually being freed from the drudgery of detailed, mechanical text manipulation so that they can focus on the bigger picture of quality assurance in information management? (Melby, 2006: 5)*

## 7.2 Applicazioni diverse per la TA

### 7.2.1. Obiettivo qualità

Soprattutto a partire dalla diffusione dei servizi gratuiti sul web (cf. 6.3), i sistemi di traduzione automatica di qualsiasi tipologia hanno iniziato ad essere utilizzati dagli utenti comuni come da enti e società per gli scopi più diversi, dai più impegnativi ai più immediati, e con un diverso grado di consapevolezza, in genere proporzionale alle conoscenze linguistiche degli utenti stessi. In alcuni casi i software di traduzione automatica vengono consultati per tradurre documenti che necessitano di essere pubblicati o utilizzati praticamente (“dissemination”; Hutchins, 2005: 7), o che per altri motivi devono garantire una qualità elevata in termini di correttezza formale e terminologica. In questo caso, come si è visto, la traduzione automatica produrrà semplicemente un “testo intermedio” (Champsaur, 2013: 23) che richiederà una revisione

---

<sup>49</sup> “Perhaps, one of the most important results of work on machine aids to translation will be an increased respect for the complexity and difficulty of creative human translation” (Melby 1981: 29).

approfondita da parte di un traduttore o editor, per correggere gli inevitabili errori e le imprecisioni del testo prodotto dal software. Tale utilizzo della TA è prevalentemente svolto all'interno delle agenzie di traduzione, al fine di contenere i costi e tentare di accelerare i tempi di produzione trasformando sempre più il lavoro dei traduttori in quello di post-editor. L'utilizzo della TA come pre-traduzione è offerto anche da diversi CAT Tools e ad essa ricorrono alle volte anche gli specialisti come supporto nella traduzione o stesura di un testo in lingua straniera. Tuttavia, ad oggi non solo i traduttori professionisti ma anche tutti coloro che hanno una certa conoscenza di una lingua straniera sono perfettamente consapevoli del fatto che i testi prodotti tramite traduzione automatica contengono sempre degli errori. Come si è visto, migliori risultati possono essere ottenuti laddove il sistema di TA sia ristretto ad un particolare ambito o linguaggio settoriale, e si basi quindi su risorse di riferimento esclusivamente limitate all'ambito relativo. In generale, anche i testi ripetitivi e con un vocabolario ridotto danno un risultato migliore e persino accettabile. Si pensi ai manuali di uso e manutenzione: se il sistema avrà accesso tramite la sua memoria o i suoi glossari alla versione precedente, potrà produrre una traduzione sufficientemente buona da essere utilizzata. In questi casi non sono rari gli accordi tra clienti e società di servizi linguistici per lo sviluppo di sistemi di traduzione automatica personalizzati e *custom-built*. Un'ulteriore possibilità di miglioramento del risultato è fornita da un pre-editing del testo da tradurre, finalizzato a semplificarne la struttura sintattica e grammaticale, accertarsi dell'assenza di errori di battitura che potrebbero causare problemi nel riconoscimento delle parole, valutare il possibile rischio di fraintendimenti e sostituire se del caso termini generici con altri più precisi. Si tratta grosso modo dello stesso processo coinvolto nella produzione dei *controlled languages*, particolarmente adatti, come si è visto, ad essere elaborati dal computer (cf. 5.2). In ogni caso, sebbene il pre-editing possa ridurre il rischio di errori nella traduzione, quando la qualità del prodotto finale deve essere garantita sarà sempre necessario sottoporre il testo anche ad un post-editing per correggere eventuali errori e portare il documento al livello degli standard espressivi e stilistici perseguiti. È quindi evidente che il tentativo di ottenere una traduzione automatica di buona qualità presupponga comunque un notevole contributo umano nella fase di preparazione del software, con la compilazione dei glossari e dei corpora di riferimento, e nella preparazione del testo con le tecniche di semplificazione e disambiguazione. L'intervento

umano è richiesto anche successivamente, per il bisogno di tenere le risorse costantemente aggiornate e mantenute, e per la necessità di compiere una revisione finale del documento. In poche parole, la traduzione automatica diventerà una HAMT, *Human-Aided Machine Translation* (Hutchins e Somers, 1992) secondo la divisione riportata in 1.1, le cui fasi sono espresse dallo schema sviluppato da Hutchins (2005), di seguito riportato:

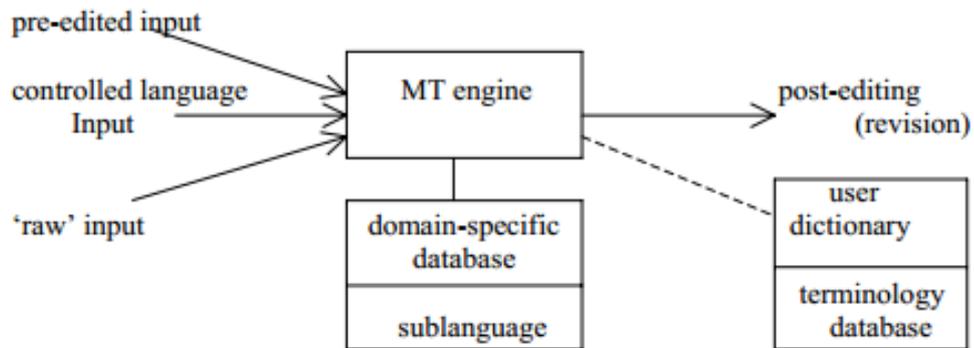


Fig. 30: Human-aided machine translation (Hutchins, 2005: 8)

Ciò implica, per le agenzie che decidono di adottare un sistema di traduzione automatica, dei costi di gestione considerevoli che rendono tale strumento poco conveniente sia dal punto di vista economico che da quello procedurale. Si tenga presente che le maggiori agenzie di traduzione lavorano con numerose combinazioni linguistiche, per ognuna delle quali devono preparare e gestire database, glossari e altri materiali per ogni ambito linguistico per il quale offrono servizi. In ultima istanza, per i testi che richiedono una qualità elevata, l'indispensabile ricorso al lavoro di linguisti specialisti fa propendere quindi nella maggior parte dei casi verso una traduzione umana assistita da computer, piuttosto che verso una traduzione automatica assistita dall'uomo. Queste considerazioni, unite a quanto si è potuto desumere dal precedente paragrafo 7.1, lasciano concludere che pensare alla traduzione automatica principalmente come uno strumento per i traduttori o alternativo ad essi sia fuorviante e lontano dalla realtà. La qualità dei testi prodotti è ancora troppo grezza e richiede pesanti interventi umani prima e dopo, e anche l'utilizzo da parte delle agenzie di traduzione è limitato a quelle occasioni in cui questo sia realmente più conveniente di una traduzione umana.

### **7.2.2. Utilizzi alternativi**

Quanto appena affermato corrisponde quindi a dire che la traduzione automatica è in realtà inutile? Come ben si sa, essa viene quotidianamente utilizzata da migliaia di persone in tutto il mondo, tanto nelle sue versioni gratuite disponibili sul web quanto sui portali appositamente sviluppati per società o uffici. Esistono infatti numerose situazioni nelle quali la TA può offrire dei vantaggi concreti. Nella maggior parte dei casi si ricorre ad essa per motivi di carattere pratico e immediato legato ad esigenze della vita quotidiana, come la traduzione di un testo di cui serve avere una rapida comprensione dei concetti essenziali (“assimilation”), la corrispondenza tra individui in lingue differenti (“interchange”), o la ricerca di informazioni in lingua straniera, per esempio sul web (“database access”). In queste circostanze, i risultati pur imperfetti del sistema assolvono sufficientemente agli scopi degli utenti, quindi non necessitano di una revisione umana e consentono di sfruttare al massimo le potenzialità dei sistemi di traduzione automatica. (Hutchins, 2005: 7; Quah, 2006: 89). Spesso i ricercatori sottovalutano l’importanza di tali utilizzi, concentrandosi sullo studio della performance dei sistemi e sul perseguimento di un miglioramento qualitativo molto difficile da ottenere. In realtà, lasciare spazio a questo aspetto vuol dire tenere conto della realtà del fenomeno e per certi versi della sua dimensione sociologica, il cui impatto è importante e significativo e dovrebbe guidare alla pari di altri aspetti lo sviluppo futuro dei sistemi di TA. Il reale utilizzo di questi strumenti è spesso molto circoscritto, finalizzato a tradurre testi ma anche singole parole o espressioni per soddisfare un bisogno pratico e sporadico di comprensione o comunicazione. Occorre tenere presente questo panorama di utilizzo a fianco di quello più strutturato e complesso già descritto, perché esso rappresenta in larga parte il vero e proprio valore aggiunto delle possibilità offerte dalla TA.

Come già accennato, la TA è ampiamente utilizzata per ottenere le informazioni essenziali contenute in documenti in lingua straniera, come pagine su internet, o per tradurre brevi stralci di testo per esigenze comunicative come nel caso della messaggistica istantanea o dello scambio di mail con riceventi stranieri. Sono questi gli utilizzi principali massicciamente svolti da utenti non specialisti per finalità pratiche e immediate, che richiedono risultati veloci e qualità non ottima, ma comunque sufficiente alla comprensione del messaggio senza fraintendimenti. Per tale scopo vengono

principalmente utilizzati i sistemi online di cui si è già parlato (cf. 6.3), accessibili gratuitamente da chiunque tramite il proprio PC.

Un esempio illustre di utilizzo alternativo dei sistemi di TA è fornito dall'Unione Europea, una delle prime istituzioni ad avvalersi della traduzione automatica in maniera estensiva (Hutchins, 2005: 9). Le varie Direzioni e Commissioni all'interno dell'UE hanno accesso ad una piattaforma intranet su cui è disponibile un sistema interno di traduzione automatica che attinge ad un vastissimo database di documenti in numerose combinazioni linguistiche. Pur adoperando tale sistema anche come pre-traduzione per i documenti ufficiali che vengono successivamente perfezionati e corretti dai traduttori prima di essere pubblicati, l'utilizzo principale che viene fatto della TA riguarda attività decisamente più pratiche e continuamente necessarie nei diversi uffici, come la rapida comprensione e produzione di corrispondenza in lingua straniera, o l'estrapolazione dei contenuti principali di documenti per valutarne l'importanza e decidere se valga la pena farli tradurre per intero ai traduttori ufficiali. Non sono quindi questi ultimi i principali fruitori di tale risorsa, quanto piuttosto dirigenti e amministratori, o il personale degli uffici, il cui bisogno di costanti consulenze linguistiche trova nei risultati della TA delle soluzioni adeguate alle finalità perseguite. La traduzione automatica può anche fornire un supporto nella stesura di testi in una lingua straniera conosciuta in maniera sufficiente ma non perfettamente padroneggiata, o consentire la comprensione di documenti che verranno esposti oralmente durante le sedute del Parlamento, al fine di preparare eventuali interrogazioni o mozioni. Il sistema di traduzione automatica dell'Unione Europea è ovviamente un sistema dall'elevato grado di affidabilità, costantemente mantenuto e basato sulla consultazione del cospicuo bagaglio di documentazione plurilingue prodotta all'interno delle istituzioni europee; può quindi assolvere in maniera soddisfacente agli scopi appena elencati, fermo restando il ricorso ai traduttori ufficiali per qualsiasi traduzione di documenti di maggiore entità e importanza.

L'utilizzo funzionale della TA viene quindi svolto tanto da utenti occasionali quanto da uffici ed enti che possono avvalersi di sistemi *custom-built*. Questa applicazione consente di trarre un reale e immediato vantaggio da strumenti anche imperfetti quali sono ancora oggi la maggior parte dei sistemi di TA. È importante dare valore a questo aspetto forse meno interessante dal punto di vista degli studi e delle sperimentazioni, ma indubbiamente più concreto e capace di mostrare la traduzione

automatica per quello che è a tutti gli effetti: oltre che un oggetto di ricerca, uno strumento al quale guardare non tanto dall'alto delle aspettative degli studi sulle intelligenze artificiali, quanto dalla realtà della sua fruizione quotidiana da parte di migliaia di utenti. Questa particolare prospettiva potrebbe anche riuscire a pacificare teorici e traduttori venendo meno le loro riserve e i timori sulla possibile equiparazione del processo traduttivo umano con quello meccanico, e sancendo una diversità di scopi, funzioni, processi e prodotti. Infine, potrebbe aprire nuovi scenari per lo studio teorico e applicato del fenomeno.

### 7.3 *Una prospettiva cognitiva*

#### **7.3.1. I modelli teorici di riferimento**

Come si è fatto per i CAT Tools, si cercherà adesso di riflettere sul fenomeno della traduzione automatica da un punto di vista particolare e poco discusso, ovvero quello cognitivo. Può sembrare strano attribuire ad un processo totalmente meccanico come la TA una dimensione cognitiva che è peculiare delle attività umane e del modo in cui il cervello umano percepisce, cataloga e riveste di significato le situazioni che sperimenta. La traduzione automatica relega al minimo l'intervento umano, viene adoperata per scopi ben precisi e non prevede, al contrario dei CAT Tools, una modifica sostanziale nelle prassi operative relative ad attività complesse come la traduzione umana, o l'introduzione di nuovi significativi elementi all'interno di procedure prestabilite. Malgrado ciò, è comunque possibile esprimere delle considerazioni che vadano al di là degli aspetti tecnici, strumentali e funzionali di tale fenomeno, e che si concentrino invece sulla percezione degli utenti circa lo strumento stesso e i testi da esso prodotti<sup>50</sup>. La percezione umana è ovviamente un elemento essenziale anche in fase di ideazione, progettazione e creazione dei sistemi. Come si è accennato all'inizio di questa Seconda Parte, uno dei presupposti dai quali si è mossa l'evoluzione dei programmi di traduzione automatica è stato quello di tentare di riprodurre in maniera meccanizzata i processi umani di comprensione e produzione linguistica. Gli studi sulle intelligenze artificiali e l'elaborazione dei linguaggi naturali hanno fornito il maggiore contributo in tal senso e permesso di guardare al meccanismo di produzione linguistica dei computer come alla

---

<sup>50</sup> Si rimanda alla lettura di Mandelbit (1993), antesignano dell'approccio cognitivo applicato alla TA

chiave per una nuova possibilità di interazione tra uomo e macchina (Moser-Mercer, 1987) basata sull'utilizzo del linguaggio umano al posto dei comandi e dei linguaggi propri dei computer. I diversi sistemi prodotti mettono l'accento su alcuni aspetti specifici della competenza linguistica, prediligendo ora la componente grammaticale (i sistemi *rule-based*), ora quella comunicativa (i sistemi *example-base*), ed in certo modo manifestando il tipo di rappresentazione del linguaggio che gli sviluppatori hanno prediletto nelle varie fasi della vita della traduzione automatica. È indubbia ad esempio la forte influenza esercitata dalla grammatica generativa di Chomsky sullo sviluppo dei sistemi *rule-based* (Melby, 2006: 2); la sua esplicitezza matematica, l'insistenza sulle formule ben formate e dal numero finito che governano un linguaggio nella sua infinità di espressioni e sulla presenza di una grammatica universale alla base della facoltà di linguaggio di ogni essere umano hanno sostenuto la convinzione di poter agevolmente riprodurre uno schema sintattico, morfologico e grammaticale e trasferirlo da una lingua ad un'altra. I diagrammi ad albero usati per l'analisi grammaticale delle frasi all'interno dei sistemi RB sono ripresi dagli indicatori sintagmatici delle grammatiche non contestuali, i *parser* utilizzati per suddividere le frasi in sintagmi a cui associare la funzione rispettiva sono prodotti facendo riferimento alla grammatica universale che codifica le proprietà grammaticali (morfologiche e sintattiche) comuni a tutte le lingue (Chomsky, 1957). Lo stesso concetto sta alla base dei sistemi *transfer-based*, che si affidano ad un'interlingua indipendente dalle lingue di partenza e di arrivo. Per gli sviluppatori di queste due tipologie di TA, allora, il linguaggio è sì capace di produrre infinite elocuzioni, ma lo fa a partire da un numero finito di regole che sono inoltre descrivibili, e che si fondano su di una base comune per tutte le lingue che rende possibile l'individuazione di costrutti simili tra due lingue e quindi il passaggio dei contenuti da essi espressi.

Dal canto loro, invece, i sistemi statistici oppongono all'idea di competenza grammaticale quella di competenza "comunicativa", esprimendo con l'estesa raccolta di testi e documenti di riferimento alla quale attingono sulla base di calcoli probabilistici e percentuali di similarità l'idea di un linguaggio troppo ricco, creativo e imprevedibile da poter essere padroneggiato a partire da un numero esiguo di regole esatte. Secondo il concetto che sta alla base dei sistemi statistici è possibile creare una frase in una data lingua per imitazione, senza necessariamente conoscere le regole grammaticali,

morfologiche e sintattiche di cui questi sistemi sono privi. Anche i modelli statistici sono debitori a Chomsky per la definizione di “esecuzione” in opposizione a “competenza (cf. 5.3). L’esecuzione è la capacità dei parlanti di utilizzare in maniera reale il linguaggio producendo frasi ed espressioni in maniera creativa. Sempre nell’ambito della sua grammatica generativa, fu Chomsky a fare notare che i bambini, senza alcuna conoscenza della grammatica formale, sono capaci di produrre enunciati per imitazione sulla base di quelli a loro noti. Se questa considerazione serviva allo studioso per avvalorare l’ipotesi di una competenza linguistica innata (*ibid.*), essa può anche essere adottata per spiegare il processo di recupero su cui si basano i sistemi statistici: maggiore il grado di somiglianza, maggiore la possibilità di ricalcare il modello linguistico desunto dall’esempio per creare una nuova frase. Maggiore sarà la raccolta di esempi a cui si potrà attingere, più facile sarà per il sistema estrapolare anche da più frasi le parti necessarie a crearne una completamente nuova. L’idea di lingua come realtà viva e apprendibile per osservazione ed imitazione è anche propria dei costruttivisti (cf. 4.1), che prediligono nella loro didattica un approccio diretto con la lingua e l’extrapolazione per metodo induttivo delle regole grammaticali (Kiraly, in Danks *et al.*, 1997).

### **7.3.2. A imitazione dell’uomo**

Le tipologie di strumenti che più si avvicinano alla competenza linguistica umana sono probabilmente i sistemi ibridi di TA (cf. 6.1), che fanno dialogare la componente grammaticale con quella comunicativa e permettono di unire i vantaggi offerti da entrambe le modalità di analisi e trasposizione dei testi. Essi rappresentano la migliore forma di imitazione del processo traduttivo umano, basato sull’interazione di competenze diverse, sia linguistiche che extralinguistiche, legate al mondo reale e alle consuetudini della comunicazione in situazioni ed ambiti specifici. I sistemi ibridi possiedono così sia una conoscenza semantica (indipendente dal contesto) fornita loro dalla loro componente RB, sia una conoscenza pragmatica (dipendente dal contesto) offerta dalla componente EB (Arnold *et al.*, 1994: 129). Il concetto di contesto che si lega a quest’ultima componente è tuttavia limitato all’osservazione della parola o della frase ricercata all’interno di una porzione più grande di testo che appunto ne contestualizzi l’utilizzo. Non si tratta di un’informazione che permette al sistema di ottenere reali informazioni sull’utilizzo di alcune espressioni legate ad esempio alla finalità del testo o al suo

contenuto, ma di un ulteriore dato che sostenga il calcolo probabilistico che sta alla base del recupero del traduttore più adeguato.

Sebbene i sistemi ibridi riescano a far superare alcune delle più comuni difficoltà di traduzione, principalmente nell'accordo tra le varie parti del discorso e nell'evitare alcuni fraintendimenti nell'utilizzo dei termini, non si può comunque attribuire a questi sistemi alcune capacità esclusive dell'uomo, come l'identificazione e applicazione delle adeguate strategie traduttive, la capacità di *problem - solving*, la disambiguazione nell'utilizzo del linguaggio in maniera peculiare e non convenzionale. Vi è sempre una incolmabile differenza di fondo tra l'intelligenza umana e quella digitale, data dalla loro natura particolare e la maniera essenzialmente diversa di percepire, assimilare ed elaborare le informazioni, che fa sì che il tentativo di riproduzione della prima da parte della seconda sia sempre e solo parziale.

### **7.3.3. Sistemi semi-automatici**

Come puntualizzato da Déjean Le Féal (1990), i diversi sistemi di TA, per quanto diversi, partono tutti da un approccio linguistico alla traduzione, che l'autrice reputa incapace di raggiungere lo scopo fondamentale della traduzione, ovvero quello di "riprodurre il contenuto cognitivo" del documento. Il motivo principale per questo insuccesso è secondo l'autrice il fatto che le lingue non esprimono la realtà nella stessa maniera, quindi non si può quasi mai tracciare una perfetta corrispondenza tra le parole della lingua di partenza e di quella di arrivo, o affidarsi all'ipotesi che frasi con un elevato numero di parole simili al loro interno esprimano concetti simili. L'autrice suggerisce una forma di traduzione semi-automatica, nella quale l'elaborazione computerizzata sia utilizzata come strumento per assistere nelle due fasi principali del processo, ovvero l'estrazione del contenuto cognitivo e la sua ri-espressione. Entrambe queste due fasi implicano due passaggi di de-verbalizzazione, la prima delle quali è il risultato dell'analisi cognitiva (non verbale) del testo di partenza, la seconda prepara invece alla riformulazione del contenuto cognitivo nel linguaggio di arrivo. Le due fasi di de-verbalizzazione richiamano i modelli *transfer-based* che ipotizzano la presenza di una rappresentazione intermedia tra le due lingue. Tuttavia, in questo modello la de-verbalizzazione non è finalizzata a superare il gap esistente tra la lingua di partenza e quella di arrivo; serve piuttosto a riformulare il pensiero dell'autore nella lingua di partenza in modo da renderlo facilmente elaborabile da parte del computer (code-

switching) con il contributo dell'uomo che fornisce al computer delle indicazioni sul senso contenuto dal testo, e ad esprimere lo stesso senso con parole della lingua di arrivo sulla base del materiale elaborato dal computer. Fondamentalmente, quindi, le due parti principali del processo traduttivo sono svolte dall'uomo, col supporto della macchina nelle fasi di scomposizione del testo di partenza e di ricerca delle migliori espressioni per la formulazione del testo di arrivo. Sebbene si possa obiettare che tale strumento non sia realmente automatico, la sua semi-automazione offre una proposta diversa di utilizzo delle risorse informatiche alla luce della grande complessità del processo traduttivo che, per essere applicato a qualsiasi tipologia e contenuto testuale, non può prescindere dal contributo umano come più volte si è puntualizzato. Il contenuto cognitivo, che è la parte più importante del testo, non può essere colto dal computer, che si ferma al livello del significativo, ovvero delle parole che compongono il testo stesso. Questo è il vero limite di tutti i sistemi di traduzione automatica a prescindere dal modello adottato o dalla combinazione linguistica interessata. L'idea di trasformare il processo di traduzione in due set di operazioni interlinguistiche guidate dagli input forniti dall'utente riproduce quella parte del processo che i sistemi automatizzati non sono in grado di svolgere con la stessa significanza cognitiva che essa ha nel processo umano, ovvero la completa de-verbalizzazione (Déjean Le Féal, 1990: 719). Il sistema risulta più simile alla traduzione umana nella tecnica, e a quella automatica nel ricorso all'elaborazione digitale dei dati per il trasferimento da una lingua all'altra. La semi-automazione che mette in dialogo uomo e computer mantiene umano tutto ciò che non può essere meccanizzato, fornendo alla macchina quelle indicazioni che permettano di sfruttare al meglio le sue potenzialità di elaborazione e produzione di dati. Come riconosciuto dall'autrice stessa:

*“it will probably never break records in translating speed nor beat the lowest cost per page translated, but should nonetheless help promote communication across language and culture barriers by providing high-standard translations that will not undermine the distinct identity and integrity of the target language” (ibid.).*

Anche Wallis (2006 e 2008) giunge a conclusioni simili, sottolineando i benefici in termini di qualità e di soddisfazione degli utenti (traduttori) di una traduzione interattiva che veda uomo e macchina in dialogo per un reciproco sostegno.

#### **7.3.4. La percezione degli utenti**

Come viene percepita la traduzione automatica come strumento, e come vengono lette le traduzioni da essa prodotte? Pur se non sono stati condotti studi specifici su questo aspetto, è comunque possibile avanzare delle ipotesi sulla base delle sperimentazioni e della letteratura prodotta relativamente ad altri aspetti, nonché tenendo in considerazione quanto detto al paragrafo 7.1. Prima di tutto, occorre anche in questo caso diversificare. Come si è visto, i traduttori hanno molto spesso un'attitudine negativa nei confronti della TA per i motivi che sono già stati presentati. Si avvicinano quindi a questo strumento e alle traduzioni da esso prodotte con diffidenza e un atteggiamento di critica, pienamente motivata dalla scarsa qualità di queste e dal target qualitativo elevato che essi hanno in mente. Gli utenti occasionali e non specialisti sono in genere più ottimisti nel ricorrere alla traduzione automatica; pur consapevoli la maggior parte delle volte dell'imperfezione dei risultati che otterranno, non sembrano essere frenati da questo aspetto, come dimostra il gran numero di utilizzazioni quotidiane dei sistemi di TA. Nel loro caso, conta più il potere avere accesso al contenuto di un testo in lingua straniera, o esprimere un concetto in maniera sufficientemente chiara per l'interlocutore straniero. Relativamente all'approccio con il testo tradotto, la prima sensazione che accomuna i diversi utenti è in generale la percezione dell'innaturalità del testo, causata dalla presenza di errori grammaticali e sintattici, nonché dalla costruzione non solo erranea, ma spesso anche illogica delle frasi. Questa sensazione si ha fondamentalmente quando si traduce verso la propria lingua madre. La traduzione verso una lingua straniera non viene mai svolta dai traduttori professionisti, ma è verosimile che essa venga svolta dagli utenti generici, soprattutto per esigenze di comunicazione o per la redazione di brevi testi. In questo caso, la minore padronanza della lingua non dà adito alla stessa sensazione di innaturalità, e fa subentrare una maggiore fiducia nel prodotto della traduzione automatica. Venendo invece all'approccio dei traduttori, qualche informazione su questo aspetto può essere desunta dagli studi sul post-editing dei testi tradotti automaticamente, che rivelano alcuni aspetti del modo in cui traduttori o revisori professionisti leggono e intervengono su questi testi (Doherty, O'Brien, Carl, 2010; Allen, Hogan, 2010; Krings, 2001). Occorre tuttavia tenere presente che, puntando alla valutazione della performance dei diversi sistemi, i dati raccolti non possono essere considerati privi di condizionamenti. In generale vale quanto detto sull'utilizzo dei match derivati dalle TM all'interno dei CAT

Tools (cf. 3.5): da un lato si può cogliere una marcata sfiducia, unita alla tendenza a iper-correggere, ovvero modificare non solo quello che è effettivamente sbagliato, ma anche quelle parti che possono essere considerate corrette e alle quali viene preferito una maniera alternativa per esprimere lo stesso contenuto. Dall'altra parte, è possibile anche il contrario, ovvero che un revisore poco accorto accetti senza verificare i termini utilizzati nella traduzione automatica, a causa di quella influenza che le tecnologie e i loro suggerimenti esercitano soprattutto sugli studenti (cf. 4.2) o sui traduttori alle prime armi. Questo può rivelarsi molto dannoso tenendo conto che la revisione dei testi prodotti automaticamente viene svolta quando si punta a risultati di alta qualità, ovvero quando la TA è utilizzata per testi di una certa importanza, solitamente specialistici, che devono essere pubblicati o utilizzati dai clienti che li hanno commissionati per il proprio lavoro o attività. Ciò trova riscontro in quanto espresso da una delle traduttrici intervistate nel questionario di cui si è parlato in 7.1, la quale ha affermato di non avvalersi della TA per pre-tradurre, per il timore di utilizzare per svista o per eccessiva fiducia la terminologia sbagliata inserita dal software. Una maggiore sicurezza viene data in tal senso dall'utilizzo della TA all'interno di un CAT Tool, tramite il quale poter verificare in tempo reale e automaticamente la *terminology consistency* e apportare dove necessario quelle modifiche che possano portare il testo agli standard di una traduzione umana.

## 7.4 *Che cosa prospetta il futuro?*

### **7.4.1. Ancora molta strada da fare**

Dopo aver guardato al passato e al presente della traduzione automatica, è ora di pensare a quale sarà il futuro di questo fenomeno secondo quelle che sono state le ultime evoluzioni in questo settore e i traguardi a cui gli studiosi accennano nella letteratura più recente. Tra i maggiori centri di ricerca che si interessano attualmente di questo settore si possono annoverare alcune università quali la Carnegie Mellon University, l'Università di Montreal, la Dublin University, la Kyoto University, ed alcune società private come AT&T, IBM, Microsoft, Xerox etc. (Quah, 2006: 152). Ancora oggi molti dei progetti che esse seguono sono in forma sperimentale o di prototipo, cosa che lascia intuire il grande potenziale di sviluppo che la traduzione automatica potrà avere nell'immediato o nel medio termine. È opinione comune tra gli studiosi di ogni categoria che, malgrado la

qualità tutt'altro che perfetta dei risultati da essa ottenuti, la traduzione automatica continuerà a diffondersi e ad essere utilizzata in sempre maggiori ambiti e con una sempre maggiore intensità in risposta al grande bisogno di comunicazione senza barriere del mercato e dell'informazione mondiale (Pym, 2013). Giocano a suo favore la maggiore convenienza in termini economici a confronto col ricorso alla traduzione umana, l'ottenimento immediato dei risultati e la sua ottimale funzione di consultazione nel caso di difficoltà e necessità linguistiche risolvibili anche con una traduzione dalla qualità mediocre (cf. 7.2). La questione qualitativa rimane comunque aperta e rappresenta l'aspetto più spinoso del fenomeno. Nell'ultimo decennio, infatti, la traduzione automatica non è migliorata significativamente dal punto di vista della qualità della traduzione prodotta, sebbene si possa apprezzare un notevole miglioramento rispetto ai primi modelli sviluppati. I progressi dal punto di vista linguistico sono stati più lenti a causa della difficoltà di elaborazione dei linguaggi naturali, legati inevitabilmente alla componente culturale e alla naturale creatività delle lingue, aspetti che i computer non sono in grado di gestire (Hutchins 2005: 21-22), come si è più volte puntualizzato. D'altro canto, sempre relativamente al miglioramento qualitativo della TA, molto si è potuto fare per i sistemi ristretti e sviluppati per uno specifico ambito, come i sistemi *custom-built* o quelli che utilizzano i linguaggi controllati, che rappresentano ad oggi la migliore prospettiva nello sviluppo di sistemi di TA qualitativamente adeguati (Graham *et al.*, 2012). Se invece ci si riferisce ai sistemi generici, la possibilità di ottenere traduzioni accurate è ancora lontana malgrado la discreta soglia di miglioramento raggiunta. Una possibilità in tal senso potrebbe essere offerta da una maggiore facilità nelle procedure di aggiornamento, modifica e miglioramento delle risorse a cui attingono i sistemi, ad esempio tramite la diretta acquisizione delle correzioni apportate dai post-editor (Thicke, 2013: 17). In questo modo, non solo i sistemi beneficerebbero di un costante intervento perfezionatore (Ambati *et al.*, 2010), ma a trarne vantaggio sarebbero anche gli stessi post-editor, che sarebbero impegnati in un lavoro più motivante e gratificante della semplice revisione di traduzioni di cattiva qualità, svolgendo un'apprezzabile mansione di QA e manutenzione dei sistemi. Ciò potrebbe anche in parte pacificare il conflittuale rapporto tra traduttori e sistemi automatici, sancendo l'indispensabilità dei primi in questa come in molte altre circostanze che richiedono l'impareggiabile apporto umano. Un altro elemento significativo per la corretta apertura dei sistemi di TA al futuro è la più diffusa

consapevolezza dei suoi limiti e della diversità delle funzioni e delle situazioni alla quale essa può essere applicata. Queste differenze devono essere chiare tanto ai ricercatori al fine di indirizzare in maniera più funzionale gli studi sul campo, quanto agli utenti, principalmente per quanto riguarda la complessità del processo traduttivo e i limiti dei sistemi automatizzati. Considerare nella giusta prospettiva i sistemi di TA rappresenterà un passo importante nell'affermazione di un utilizzo maturo e critico degli stessi da parte di un'utenza variegata che ricorrerà a sistemi specifici per specifiche esigenze, consapevole del tipo di risultato che otterrà.

Contrariamente all'aspetto qualitativo delle traduzioni prodotte, un aspetto che ha goduto di un notevole miglioramento è stato invece quello relativo alla struttura dei diversi sistemi di TA, molto più adattabili, rapidi, espandibili con risorse di riferimento potenzialmente illimitate. Le accresciute caratteristiche tecniche di questi sistemi e le loro prestazioni avanzate sono un segnale della grande sperimentazione in corso, con il lancio sul mercato di versioni nuove o aggiornate che consentano una maggiore velocità di elaborazione, il ricorso ad algoritmi sempre più sofisticati per l'identificazione dei traducanti più adatti alle circostanze, tecniche di allineamento che consentono di identificare all'interno dei testi affiancati la corrispondenza di termini e frasi tra il testo di partenza e quello di arrivo (Quah, 2006: 153). Inoltre, mentre alcune tipologie di sistemi come i *rule-based* hanno già da qualche tempo iniziato a perdere terreno, si stanno affermando sempre più delle particolari tipologie di sistemi, ritenute più promettenti in termini di flessibilità e qualità dell'output, quindi più promettenti per sviluppi futuri. Si tratta dei modelli ibridi (cf. 6.1) e dei sistemi interattivi o semi-automatici; i primi si sono rivelati i migliori come qualità delle traduzioni prodotte in maniera completamente automatica, i secondo invece traggono vantaggio dalla loro semi-automazione che contempla il contributo umano in corso d'opera, non solamente in fase di pre- o post-editing (*ibid.*). I sistemi interattivi consentono quindi all'uomo di avere pieno controllo dell'intero processo traduttivo, senza per questo ridurre il vantaggio operativo ottenibile dal ricorso all'automazione dei processi (come si è visto in 7.3).

#### **7.4.2. Le lingue minoritarie**

Al di là dell'aspetto qualitativo, numerose sfide si presentano per il futuro della traduzione automatica. Vengono ad esempio sempre più richiesti sistemi capaci di tradurre da e verso le lingue minoritarie o locali, in particolare quelle asiatiche e africane,

solitamente trascurate dai sistemi commerciali per via della poca convenienza nell'investire in combinazioni linguistiche poco richieste (Somers, 2003). Questo trend è comunque in controtendenza già da qualche tempo. Sono già stati compiuti dei passi avanti nell'inclusione di un numero sempre maggiore di lingue – basti pensare alla grande schermata di lingue disponibili sugli strumenti di traduzione automatica offerti da Google e Bing, i principali attualmente disponibili sul web. Di recente si è anche potuto notare nella produzione di nuovi sistemi un passaggio dalle combinazioni linguistiche tradizionali ad altre nuove ed insolite. Specialmente l'espansione dell'Unione Europea ha rappresentato un punto di partenza per la crescente ricerca in quest'area, per rispondere alla richiesta dei parlanti nelle 23 lingue nazionali che essa ingloba di sentirsi maggiormente partecipi e percepire le istituzioni europee più vicine tramite la produzione dei documenti ufficiali nelle varie lingue locali (Stein, 2013: 12). Maggiori sistemi dovranno essere sviluppati in futuro per le lingue minoritarie o rare, la cui conservazione e valorizzazione è oggi presa in grande considerazione nel tentativo di contrastare l'appiattimento culturale che una globalizzazione portata all'estremo potrebbe causare. La possibilità di usufruire di strumenti per la traduzione automatica da e verso queste lingue può aprire inoltre nuove opportunità di sviluppo alle popolazioni dei paesi in cui tali lingue sono adoperate (Froeliger e Ladmiral, 2010: 618). Le difficoltà in tal senso non sono soltanto economiche, legate al cospicuo investimento necessario per la formulazione di nuovi sistemi, ma sono anche legate spesso alla ridotta quantità di documenti di riferimento disponibili con i quali istruire i sistemi stessi. Una soluzione a questo problema può essere rappresentata dalla collaborazione spontanea degli utenti interessati a valorizzare queste risorse, come visto in 6.3, o il ricorso a traduttori madrelingua per la produzione di testi nelle lingue interessate come suggerito da Winkler *et al.* (2013).

### **7.4.3 Altre sfide**

Altre caratteristiche che vengono richieste ai sistemi di TA sono una maggiore interattività e la possibilità di integrare in un unico sistema strumenti commerciali (ottimali per utilizzi mirati e più precisi) e strumenti per scopi generici. Un particolare tipo di interazione che può aprire nuove opportunità è quello con i Web semantico, definito da Berners-Lee e Hendler, i suoi creatori, come “*an extension of the current Web in which information is given well-defined meaning, enabling computers and people to work in better conditions*” (cit. in Quah, 2006: 167). Le potenzialità del Web semantico

che la TA potrebbe sfruttare riguardano fundamentalmente il recupero di informazioni e documenti da utilizzare come materiale di riferimento, facendo della rete un infinito corpus in continua espansione, consultabile dai sistemi di traduzione automatica sia *example-based* che *knowledge-based*, grazie all'introduzione di informazioni semantiche nei testi, che consentano la loro estrazione sulla base dell'argomento o delle informazioni contenute in essi. L'accesso al contenuto consentirà di individuare le eventuali traduzioni in maniera più accurata e meno suscettibile a possibili ambiguità.

Il futuro potrebbe infine portare sistemi di traduzione automatica capaci di elaborare input vocali con l'ausilio della *speech technology*, sebbene essi appaiano ancora oggi progetti molto avveniristici. Non si tratterà solo di dispositivi capaci di convertire un input vocale in testo e viceversa (strumenti di questo genere esistono già sul mercato), ma anche di produrre una traduzione di un input vocale sotto forma di testo (*speech-to-text translation*) o di audio (*speech-to-speech translation*), aprendo la strada a quella che potrebbe essere definita come un'interpretazione automatica (Quah, 2006: 158). Sebbene le macchine capaci di comprendere un certo numero di comandi vocali si stiano diffondendo in numerosi settori, sistemi di questo genere richiederebbero una molto complessa mappatura di tutte le parole di almeno due lingue, fino ad arrivare allo sviluppo di sistemi per il riconoscimento vocale multilingue. Le difficoltà per lo sviluppo di sistemi simili sono evidenti, legate tanto alle variazioni di pronuncia e alle versioni locali di una lingua, quanto alla complessità del processo stesso di decodifica, trasposizione e ricodifica. Ad oggi esistono soltanto dispositivi ideati principalmente per il turismo, contenenti le frasi e le espressioni di uso più comune. Sono state sviluppate anche numerose applicazioni per smartphone e dispositivi mobili, in grado di recepire e tradurre correttamente frasi semplici e brevi, ma dalle prestazioni nettamente inferiori per frasi più lunghe e complesse. Il loro utilizzo è limitato ma comunque promettente, considerato che molte delle sperimentazioni a riguardo sono ancora in corso.

## 7.5 Conclusioni

Alla fine di questa non esaustiva trattazione, che si è sforzata di presentare gli strumenti di traduzione automatica da un punto di vista critico e il più possibile ampio per includere le diverse sfaccettature del fenomeno, è giunto il momento di trarre alcune

conclusioni che sintetizzino gli aspetti più importanti ai quali si è prestata attenzione nel corso di questi tre capitoli:

1. prima di tutto, la traduzione automatica ha una storia relativamente lunga. Non è, come molti credono, una realtà emersa negli ultimi dieci anni. Come tale, essa ha già alle spalle numerosi studi, sperimentazioni, miglioramenti e ostacoli superati. Tuttavia, grazie alla continua evoluzione delle tecnologie e all'avanzamento degli studi sull'elaborazione del linguaggio naturale e sulle intelligenze artificiali, la TA è anche una realtà proiettata nel futuro, che dovrà ancora superare numerosi limiti e migliorare su diversi fronti, specialmente quello qualitativo. Inoltre la richiesta di dispositivi per la traduzione automatica è in costante crescita, e si prevedono sempre maggiori campi di applicazione soprattutto sul web, ormai divenuto il luogo principe degli scambi e delle comunicazioni globali (cf. cap. 5 e par. 7.4);

2. esistono ad oggi numerose tipologie di sistemi di TA. Prima di tutto, esistono strumenti gratuiti disponibili sul web, come i celebri Google Translate o Bing Translator, e strumenti a pagamento offerti da agenzie di servizi linguistici, che hanno il vantaggio di poter essere personalizzabili e garantire al cliente una maggiore precisione dei risultati prodotti. Entrambe queste categorie di strumenti ricorrono a modelli diversi di traduzione. Sono ancora disponibili alcuni sistemi *rule-based*, che rappresentano i modelli più antichi anche noti come “modelli classici”. I sistemi più diffusi sono comunque ad oggi quelli statistici o *example-based*, che sembrano fornire risultati migliori in termini di comprensibilità generale del testo. Una tendenza sempre più comune è quella di utilizzare le due modalità in maniera congiunta all'interno di sistemi ibridi. Allo stesso modo, è molto frequente l'utilizzo della TA all'interno dei CAT Tools come pre-traduzione per quei segmenti che non trovano alcuna corrispondenza (match) all'interno della memoria di traduzione (cf. cap. 5 e 6);

3. la traduzione automatica può essere utilizzata per scopi diversi, il principale dei quali non è, come si può erroneamente pensare, la traduzione di testi ampi finalizzati alla pubblicazione o diffusione (applicazione per la quale ad oggi la TA è ancora incapace di fornire risultati che soddisfino le aspettative degli utenti). Quotidianamente i sistemi di traduzione automatica vengono utilizzati da migliaia di utenti in tutto il mondo per risolvere problemi di comprensione o comunicazione per i quali non è necessario uno standard qualitativo elevato delle traduzioni prodotte. Per questo tipo di applicazione la

traduzione automatica appare già come un sistema ottimale che consente il più delle volte di raggiungere gli scopi perseguiti; nel caso in cui invece essa venga utilizzata per tradurre documenti di maggiore complessità e importanza, l'intervento di un revisore umano sarà comunque necessario per apportare le necessarie correzioni. Risultati migliori sono ottenuti dai sistemi dedicati ad ambiti ristretti e specifici, o da quelli che elaborano i linguaggi controllati (cf. par. 7.2; 5.3 etc.);

4. dai precedenti due punti è chiaro che quando si parla di traduzione automatica non si fa riferimento ad un fenomeno omogeneo e ristretto, ma ad un panorama di strumenti, applicazioni, scenari e prospettive molto ampio e in continua crescita. Esistono situazioni in cui la traduzione automatica può essere utilizzata ricavandone una reale utilità, altre dove essa deve essere accompagnata da altri strumenti o supportata dall'intervento dell'uomo, altre ancora in cui essa non è capace di fornire alcun aiuto reale e conviene rivolgersi direttamente a un professionista. La traduzione automatica non è una nemica dei traduttori professionisti, perché i loro campi di applicazione sono diversi. Il fatto che diverse agenzie mascherino veri e propri lavori di traduzione sotto l'etichetta di un post-editing di traduzioni automatiche di cattiva qualità è uno dei principali motivi per i quali i traduttori hanno un rapporto di sfiducia e rifiuto nei confronti di questa tecnologia. Tuttavia, il timore che un giorno la TA scalzerà completamente i professionisti non trova corrispondenza nella realtà oggi e probabilmente non la troverà neanche in futuro;

5. dal punto di vista cognitivo, trattandosi di un'attività interamente automatizzata, essa non ha grandi ricadute sull'uomo. Piuttosto è vero il contrario, cioè che la percezione degli studiosi e degli sviluppatori dei software relativa al processo traduttivo e al linguaggio è stata la base sulla quale i diversi sistemi hanno cercato di riprodurre in qualche modo l'abilità tutta umana di comprendere e riprodurre una o più lingue. Il rapporto degli utenti con lo strumento "traduzione automatica" è diverso se si prendono in considerazione i traduttori professionisti – generalmente avversi ad esso – e gli utenti generici ed occasionali che pur consapevoli dell'ampio margine di errore a cui possono andare incontro, utilizzano questa risorsa con discreta fiducia. Gli studi sul post-editing dei testi tradotti automaticamente rivelano alcuni aspetti del modo in cui questi testi vengono letti, prevalentemente da traduttori o revisori professionisti, ma puntando alla valutazione della performance dei diversi sistemi i dati raccolti non possono essere considerati privi di condizionamenti;

6. occorre infine ricordare che la traduzione è uno strumento, non uno scopo. È un metodo per facilitare la comunicazione interlinguistica. Tutti gli strumenti digitali di sussidio alla traduzione, automatici e non, rimangono appunto strumenti, e come tali occorre valutare la loro efficienza, facilità di utilizzo, versatilità e risposta alle esigenze dell'utente oltre e forse più che la qualità della traduzione prodotta, consapevoli del fatto che essa non sarà mai perfetta, ma che anche la sua imperfezione può tornare a beneficio di traduttori e utenti comuni, fornendo un testo "grezzo" che dovrà essere – ove più, ove meno – riveduto dall'uomo e adattato alle esigenze di utilizzo reale. È necessario riconoscere che la traduzione automatica si è ormai affermata come strumento ma anche come fenomeno sociale, che ha aperto a molte persone degli scenari e delle possibilità nuove nella fruizione e nella produzione di testi in lingue non conosciute, malgrado la qualità non ottimale degli output di questi sistemi. Occorre superare antiche diatribe e opposizioni, e imparare a distinguere le situazioni e gli scopi per i quali la traduzione automatica, alla pari di qualsiasi altra risorsa, può essere d'aiuto concretamente. Ciò consentirà di valorizzare questo strumento nella maniera più giusta, riconoscendone i meriti innegabili al di là dei limiti che esso ha ancora e che probabilmente avrà anche in futuro.

## Conclusioni generali

La trattazione che volge al termine ha inteso offrire una presentazione il più possibile completa delle diverse tecnologie applicate alla traduzione, sforzandosi di mantenere l'attenzione sul destinatario ultimo per cui esse sono state prodotte, ovvero l'uomo. La traduzione è essenzialmente un'attività umana, e malgrado i diversi aspetti che essa può assumere, le diverse esigenze che la rendono necessaria, i molteplici scopi che essa può perseguire e le varie modalità tramite le quali essa può essere prodotta, ciò che resta immutata è la sua natura di atto comunicativo umano, particolarissimo per l'insieme di conoscenze che esso presuppone, per le dinamiche che attiva e per il ruolo di mediatore che il traduttore viene a svolgere in esso <sup>51</sup>.

Ma perché parlare di tecnologia se l'intento è quello di studiare la traduzione come fenomeno e processo umano? A questa domanda si potrebbe rispondere con un'altra semplice domanda: perché no? L'opinione comune in fatto di *translation tools* è che essi poco abbiano da dire alla ricerca di tipo linguistico e allo studio della traduzione come disciplina. Essi vengono prevalentemente studiati da settori di ricerca più marcatamente tecnici e applicati, e con approcci empirici atti a valutarne l'efficienza, la precisione, l'usabilità. Anche all'interno degli studi linguistici, la loro trattazione è spesso relegata al contributo materiale apportato al processo traduttivo, e al massimo a qualche nota tecnica sulla necessità di includere il loro studio all'interno della didattica della traduzione.

In realtà, come suggerito da Pym (2011b), è importante che la ricerca umanistica si riappropri di una sua voce in capitolo in fatto di tecnologie *applicate alla traduzione*, essendo l'unica in grado di dire qualcosa sui processi propri della traduzione *applicati alle tecnologie*, e inizi a sviluppare un approccio critico nei confronti di questo fenomeno, come ad onore del vero un certo numero, seppur limitato, di studiosi fa già da più di un decennio.

---

<sup>51</sup> Cf. Neubert in Danks et al., 1997

Volendo adottare in questa sede una prospettiva dichiaratamente umanistica, si è deciso di considerare in maniera prioritaria quegli aspetti afferenti alla così detta traduttologia cognitiva, che avrebbero potuto contribuire a svelare alcuni aspetti dell'interazione tra l'uomo e la macchina, tanto relativamente ai sistemi *computer – aided* quanto a quelli completamente automatici che parrebbero a primo acchito escludere a priori qualsiasi contributo umano al processo.

L'adozione dei CAT Tools come strumenti di gestione del processo traduzione, che essi accompagnano in tutte le sue fasi accelerando e semplificando alcune delle attività più meccaniche che esso include, ha causato delle ricadute dirette – sia positive che negative – sulle prassi operative consolidate del traduttore, sul suo *modus operandi*, sulle nuove istanze professionali che egli è chiamato ad affrontare come il lavoro in team e lo scambio di risorse. L'impatto maggiore, tuttavia, è stato esercitato su aspetti più profondi e significativi come la sua percezione del testo da tradurre, del lavoro che svolge e del proprio ruolo all'interno del processo traduttivo.

Anche la traduzione automatica, considerata erroneamente da molti come alternativa alla traduzione umana, è uno strumento che può rivelare maggiori aspetti di sé grazie al confronto con quei processi e quelle capacità umane il cui tentativo di riproduzione automatizzata ha prodotto i diversi modelli oggi esistenti. Inoltre, tenendo a mente la percezione che i suoi utenti hanno dei processi ad essa afferenti e dei documenti da essa prodotti – percezione sostanzialmente diversa per traduttori e utenti comuni – è possibile comprendere meglio le finalità alle quali essa può dare una migliore risposta, e in ultima istanza distinguere tra gli ambiti di applicazione propri della traduzione umana e gli ambiti in cui la traduzione automatica può essere applicata in maniera funzionale.

Le conclusioni alle quali si è pervenuti a partire dagli studi e dalle sperimentazioni prese in esame confermano la validità dell'approccio adottato, che promette di fornire elementi utili al miglioramento dei sistemi stessi sulla base della maggiore conoscenza della percezione che di essi hanno gli utenti, e di consentire a questi ultimi di relazionarsi in maniera più consapevole e critica a questi strumenti. Sono tuttavia conclusioni parziali. Al di là del questionario d'indagine di cui si sono riportati i risultati, e del riferimento alle esperienze didattiche pregresse agli studi dottorali che sono sfociati nel presente lavoro, la tesi manca evidentemente di una parte sperimentale vera e propria che avrebbe potuto fornire dati ed evidenze a sostegno di quanto ipotizzato circa le ricadute dell'uso dei CAT

Tools sui comportamenti cognitivi dei traduttori. Ciò può essere in parte giustificato dalla difficoltà oggettiva di osservazione di questi aspetti, per i quali si è fatto ricorso negli studi citati a strumentazioni tecnologiche specifiche e costose come i rilevatori dei movimenti oculari e del comportamento al computer (pause, uso di mouse e tastiera ecc.). Esistono tuttavia metodologie di analisi più semplici, come i *Think-Aloud Protocols* o le altre forme di verbalizzazione retrospettiva, tramite i quali è possibile ottenere informazioni parziali ma pur sempre utili a determinare ad esempio il grado di consapevolezza e controllo del processo traduttivo, e i meccanismi decisionali messi in atto. Anche l'esperienza didattica, che per vicissitudini interne all'ateneo palermitano si è dovuta interrompere essendo venuti a mancare i corsi specificamente dedicati all'applicazione delle tecnologie in ambito linguistico, non si è potuta avvalere delle nuove conoscenze e conclusioni alle quali si è pervenuti nel corso degli anni in cui la presente tesi è stata sviluppata. Anche l'ambito della traduzione automatica, che ha rappresentato per chi scrive un campo di indagine nuovo, avrebbe potuto beneficiare di un maggiore approfondimento delle architetture ibride, più vicine come si è visto al modello di competenza traduttiva umana, e delle modalità di utilizzo interattivo dei sistemi semi-automatici che introducendo il contributo umano all'interno del processo meccanico aprono prospettive nuove per il futuro della Machine Translation.

A conclusione di questa tesi, ci si augura di aver rassicurato (e magari anche incuriosito) quanti hanno sempre guardato con ostilità o scetticismo alle tecnologie applicate al linguaggio, e di avere allo stesso tempo messo in guardia i più entusiasti sulle ricadute negative che un uso eccessivamente fiducioso di questi strumenti può comportare. Si spera anche che la riflessione critica compiuta abbia in qualche modo contribuito a fornire prospettive stimolanti sullo studio della traduzione come fenomeno cognitivo, e sulle tecnologie sviluppate per venire in aiuto ad esso. Si spera inoltre di aver contribuito al superamento delle inutili e anacronistiche dicotomie che oppongono favorevoli e contrari, tradizionalisti e innovatori, studiosi teorici e traduttori per mestiere. È impensabile ad oggi non riconoscere la realtà della traduzione come settore del mercato, ed il fatto che essa possa contare su strumenti atti ad incrementare la produttività e il perseguimento di determinati standard qualitativi. Per quanto si possano nutrire perplessità su di essi, questi strumenti rappresentano oggi la chiave di accesso all'industria della traduzione, ed il loro utilizzo è spesso una scelta obbligata per chi

lavora in questo ambito. Le innegabili motivazioni economiche soggiacenti all'affermazione delle tecnologie non sono tuttavia in antitesi con la possibilità di compiere su di esse degli studi teorici ed empirici che consentano di arricchire la discussione a loro riguardo con contributi che oppongano ai meri interessi di produzione e competitività gli interessi altrettanto importanti degli utenti primari di questi strumenti, ovvero i traduttori, e ancor più gli interessi della traduzione stessa quale fenomeno linguistico complesso da valorizzare, comprendere e tutelare.

## Appendice

### QUESTIONARIO DI INDAGINE

(si prega di rispondere nella maniera più esaustiva possibile)

- 1) Età dell'intervistato:
- 2) Da quanti anni ti occupi di traduzione?
- 3) Qual è il principale settore per il quale traduci?
- 4) Utilizzi abitualmente un CAT Tool? Se sì, quale?
- 5) Quali sono a tuo avviso i maggiori benefici dell'uso di un CAT Tool?
- 6) Quali altri tipi di risorse e strumenti utilizzi per il tuo lavoro?
- 7) Ti capita di ricorrere agli strumenti di traduzione automatica, per esempio per una pre-traduzione del testo?
- 8) Che cosa pensi in generale degli strumenti per la traduzione automatica?
- 9) Qual è la tua esperienza con i CAT Tools e con la traduzione automatica?
- 10) Ti è capitato di avere richiesto da un'agenzia per cui hai lavorato di utilizzare un particolare tool o di ricorrere alla traduzione automatica?
- 11) Come utilizzi internet per aiutarti nel tuo lavoro?

## Bibliografia

- AHLSÉN, E. 2006. *Introduction to Neurolinguistics*. Amsterdam/Philadelphia, John Benjamins.
- AIKAWA, T.; L. Schwartz; R. King; M. Corston-Oliver; C. Lozano. 2007. "Impact of controlled language on translation quality and post-editing in a statistical machine translation environment". In *Proceedings of the MT Summit XI*: 1-7.
- ALCINA, A. 2008. "Translation technologies scope, tools and resources". In *Target*, 20 (1): 79-102.
- ALLEN, J. e C. Hogan. 2000. "Towards the Development of a Post-editing Module for Raw Machine Translation Output: A Controlled Language Perspective". In *Proceeding of the Third International Controlled Language Applications CLAW 2000*, Seattle: 62-71.
- ALVES, F. e A. Hurtado Albir. 2010. "Cognitive approaches". In *Handbook of Translation Studies*, Vol. 1: 28-35. Amsterdam/Philadelphia, John Benjamins.
- ALVES, F. e D. Couto Vale. 2009. "Probing the Unit of Translation in Time: Aspects of the Design and Development of a Web Application for Storing, Annotating and Querying Translatin Process Data". In *Across Languages and Cultures*, 10 (2): 251 - 273
- ALVES, F. e T. Liparini Campos. 2009. "Translation technology in time: investigating the impact of translation memory systems and time pressure on types of internal and external support". In S. Göpferick, A. Jakobsen, I. Mees (eds.) *Behind the Mind. Methods, Models and Results in Translation Process Research*. Copenhagen Studies in Language, 37: 191-218.
- AMBATI, V.; S. Vogel; J. G. Carbonell. 2010. "Active Learning and Crowd-Sourcing for Machine Translation". In *LREC* (1).

- AMIGO EXTREMERA, J. 2013. "Horizontes lejanos: traducción (inter)cultural y traductología cognitiva". In *Proceedings of the 31st AESLA International Conference "Communication, Cognition and Cybernetics"*.
- ANDREYEWISKY, A. 1981. "Translation: Aids, Robots and Automation". In *Meta*, 26 (1): 57-66. Disponibile all'indirizzo <http://id.erudit.org/iderudit/001888ar>.
- ARNOLD, D. 2000. "Why translation is difficult for computers." In *Computers and Translation: a handbook for translators*. Amsterdam/Philadelphia, John Benjamins.
- ARNOLD, D. e Y. Berglund. 1998. "WWW access to corpora: a tool for teaching and learning about corpora". In *TALC-98 (Third International Conference on Teaching and Language Corpora)*, Keeble College, Oxford. Humanities Computing Unit, University of Oxford.
- ARNOLD, D.; L. Balkan; S. Meijer; R. L. Humphreys; L. Sadler. 1994. *Machine Translation. An Introductory Guide*. Londra, NCC Blackwell.
- ARNON, I. e N. Snider. 2010. "More than Words: Frequency Effects for Multi-Word Phrases". In *Elsevier - Journal of Memory and Language*, Vol. 62: 67-82.
- AUBIN, M. C. 2010. "La localisation en 2009: la fin d'un rêve". In *Meta*, 22 (2): 119-135. Disponibile all'indirizzo <http://id.erudit.org/iderudit/1009118ar>.
- AUSTERMÜL, F. 2010. "A collaborative Approach to the Teaching of Terminology Management". In *T21N – Translation in Transition*. Vol. 2010 (09).
- AUSTERMÜL, F. 2001. *Electronic Tools for Translators. Translation Practices Explained*. Manchester, St. Jerome Publishing.
- BADDELEY, A. D. 2000. "The Episodic Buffer: A New Component of Working Memory?" In *Trends in Cognitive Sciences*, 4 (11), Elsevier Science Ltd.
- BADDELEY, A. D. 1992. "Working Memory". In *Science*, 255: 556 - 559
- BADDELEY, A. D. e G. Hitch. 1974. "Working memory". In *Psychology of learning and motivation*, 8: 47-89.
- BALKAN, L. 1992. "Translation Tools". In *Meta*, 37 (3): 408-420. Disponibile all'indirizzo <http://id.erudit.org/iderudit/001954ar>.
- BALLARD, M. 2010. "Unit of Translation". In *Handbook of Translation Studies*, Vol. 1: 437-440. Amsterdam/Philadelphia, John Benjamins.

- BARRACHINA, S. et al. 2009. "Statistical Approaches to Computer-Assisted Translation". In *Computational Linguistics* 35 (1): 3-28.
- BASSNETT, S. 1991. *Translation Studies*. Londra/New York, Routledge.
- BAZRAFESHAN, M. e D. Gildea. 2013. "Semantic Roles for String to Tree Machine Translation". In *ACL* (2): 419-423.
- BERNARDINI, S. 1999. "Using Think-Aloud Protocols to investigate the translation process: methodological aspects". In *RCEAL Working papers in English and applied linguistics* 6, J. N. Williams (ed.), University of Cambridge: 179-199.
- BERNARDINI, S. 2001. "Think-aloud Protocols in Translation Research: Achievements, Limits, Future Prospects". In *Target*, 13 (2): 241-263.
- BIAU GIL, J. R. e A. Pym. 2006. "Technology and Translation (A Pedagogical Overview)". In A. Pym, A. Perestrenko, B. Starink (eds.), *Translation Technology and its Teaching*. Intercultural Studies Group, Universitat Rovira i Virgili; Tarragona (Spagna): 5 – 19.
- BIÇICI, E. e M. Dymetman. 2008. "Dynamic Translation Memory: Using Statistical Machine Translation to Improve Translation Memory Fuzzy Matches". In *Lecture Notes in Computer Science*, 4919: 454-465.
- BLANCAFORT, H. et al. 2010. "TTC: Terminology Extraction, Translation Tools and Comparable Corpora". In *14th EURALEX International Congress, Leeuwarden/Ljouwert (Netherlands)*: 263-268.
- BOLANOS-MEDINA, A. e J. Isern Gonzáles. 2012. "Análisis de las actitudes de los estudiantes hacia las herramientas informáticas de traducción asistida". In *SENDEBAR*, Num. 23: 275-300
- BONONNO, R. 2000. "Terminology for Translators – An Implementation of ISO 12620". In *Meta* 45 (4): 646-669. Disponibile all'indirizzo <http://id.erudit.org/iderudit/002101ar>.
- BOWKER, L, e D. Fisher. 2010. "Computer-Aided Translation". In *Handbook of Translation Studies*, Vol. 1: 60-65. Amsterdam/Philadelphia, John Benjamins.
- BOWKER, L. e M. Barlow. 2004. "Bilingual Concordancers and Translation Memories: a Comparative Evaluation. In Association for Computational Linguistics, *Proceedings of*

*the Second International Workshop on Language Resources for Translation Work, Research and Training: 70-79.*

BOWKER, L. 2002. *Computer-Aided Translation Technology. A Practical Introduction*. Ottawa: University of Ottawa Press.

BOWKER, L. e J. Pearson. 2002. *Working with Specialized Language: A Practical Guide to Using Corpora*. London: Routledge.

BOWKER, L. 1998. "Using Specialized Monolingual Native-Language Corpora as a Translation Resource: A Pilot Study". In *Meta*. 43 (4): 631-651. Disponibile all'indirizzo <http://id.erudit.org/iderudit/002134ar>.

BRINKMANN, K. 1981. "Machine Aids to Translation". In *Meta*, 26 (1): 67-75. Disponibile all'indirizzo <http://id.erudit.org/iderudit/002145ar>.

BROWN, P. et al. 1990. "A Statistical Approach to Language Translation" In *Proceedings of the 12<sup>th</sup> conference on Computational linguistics* – Vol. 1: 71-76.

BRUNETTE, L. e M. Charron. 2006. "Langue, traduction et mondialisation: interactions d'hier, interactions d'aujourd'hui". In *Meta*, 51 (4): 739-743. Disponibile all'indirizzo <http://id.erudit.org/iderudit/014338ar>.

CALVI, M. V. 2003. "La traduzione nell'insegnamento della lingua e nello studio dei linguaggi specialistici". In *Tradurre dallo spagnolo – Giornata di studi*, Milano 28 febbraio 2003.

CALZADA PEREZ, M. 2005. "Applying Translation Theory in Teaching". In *New Voices in Translation Studies* (1): 1-11.

CALZOLARI, N. e A. Lenci. 2004. "Linguistica computazionale – Strumenti e risorse per il trattamento automatico della lingua", in *Mondo Digitale – Rassegna critica del settore ICT*, giugno 2004. AICA Milano.

CAMPBELL, S. 1991. "Towards a Model of Translation Competence". In *META* 36 (2-3): 329-343. Disponibile all'indirizzo <http://id.erudit.org/iderudit/002190ar>.

CARTONI, B. e T. Meyer. 2012 "Extracting Directional and Comparable Corpora from a Multilingual Corpus for Translation Studies". In *Proceedings of the eighth international conference on Language Resources and Evaluation (LREC)*.

CHAMPSAUR, C. 2013. "La traduction automatique: un outil pour les traducteurs?" In *Journal of Specialised Translation*, Num. 19 : 19-27.

- CHANG, B., K. Byeong Kwu, e Y. Shiwen. 2006. "Developing CAT Tools for Translating Chinese Monographs". Disponibile all'indirizzo <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.126.6565>.
- CHIANG, D. 2007. "A Hierarchical Phrase-Based Translation Representations". In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*: 1373–1383. Edimburgo.
- CHIARI, I. 2008. "La chiave probabilistica delle lingue: teoria linguistica e applicazioni computazionali". In L. Fulci e E. Sciubba (a cura di), *Linguaggio, mente e società*, Roma, EuRoma - La Goliardica: 55-79.
- CHOMSKY, N. 1957. *Syntactic Structures*. The Hague.
- CHRISTENSEN, T. P. 2011. "Studies on the Mental Processes in Translation Memory-Assisted Translation – The State of the Art". In *Trans-kom*. 4 (2): 137-160.
- CHRISTENSEN, T. P. e A. Schjoldager. 2011. "The Impact of Translation-Memory (TM) Technology on Cognitive Processes: Student – Translators' Retrospective Comments in an Online Questionnaire". In *Human-Machine Interaction in Translation: Proceedings of the 8th International NLPCS Workshop*, 41: 119-130.
- CHRISTENSEN, T. P. e A. Schjoldager. 2010. "Translation-memory (TM) research: What do we know and how do we know it". In *Hermes–Journal of Language and Communication Studies*, 44: 89-101.
- CIOLA, B. 2008. "La terminologia orientata alla traduzione: nuove realtà, nuovi approcci". In *Atti del convegno La Traduzione. Professione e formazione*. Università di Padova. Disponibile all'indirizzo <http://www.cat-trainer.eu/files/terminologia-PD.pdf>
- COLLOMBAT, I. 2006. "General Knowledge: A Basic Translation Problem Solving Tool". In *Translation Studies in the New Millennium - An International Journal of Translation and Interpreting*. Vol. 4.
- CORNESS, P. 1987. "The ALPS Computer-Assisted Translation System in an Academic Environment". In Picken, C. (ed.) *Translating and the Computer 7*. Londra, Aslib: 118-127.
- COUGHLIN, J. 1990. "MT or CAT: Criteria for the Independent Translator to Select a Computerized System". In *Meta* 35 (2): 455-457. Disponibile all'indirizzo <http://id.erudit.org/iderudit/002415ar>.

- DALOISO, M. 2009. *I fondamenti neuropsicologici dell'educazione linguistica*. Venezia, Libreria Editrice Cafoscarina.
- DANCETTE, J. e S. Halimi. 2005. "La représentation des connaissances ; son apport à l'étude du processus de traduction". In *Meta* 50 (2): 548-559. Disponibile all'indirizzo <http://id.erudit.org/iderudit/010999ar>.
- DANKS, J.; G. Shreve; S. Fountain; M. McBeath (ed). 1997. *Cognitive Processes in Translation and Interpreting*, Sage Publications – Applied Psychology Vol.3.
- DASCAL, M. e I. E. Dror. 2005. "The Impact of Cognitive Technologies". In *Pragmatics and Cognition*, 13: 451-457.
- DEE LUKAS, D. e J. H. Larkin. 1995. "Learning from Electronic Texts: Effects of Interactive Overviews for Information Access". In *Cognition and Instruction*, 13 (3): 431-468.
- DEJEAN LE FEAL, K. 1990. "A different Approach to Machine Translation". In *Meta* 35 (4): 710-719.
- DE PEDRO RICOY, R. 2007. "Internationalization vs. Localization: The Translation of Videogame Advertising". In *Meta* 52 (2): 260-275. Disponibile all'indirizzo <http://id.erudit.org/iderudit/016069ar>.
- DESILETS, A. *et al.* 2009. "How Translators Use Tools and Resources to Resolve Translation Problems : An Ethnographic Study". In *Proceedings of Beyond Translation Memories: New Tools for Translators MT Summit XII*
- DILLINGER, M. e A. Lommel. 2004. *LISA Best Practices Guide: Implementing Machine Translation*. Ginevra, Localization Industry Standards Association.
- DIJKSTRA, T. *et al.* 2010. "How cross-language similarity and task demands affect cognate recognition". In *Elsevier – Journal of Memory and Language*, Vol. 62: 284-301.
- DI MARTINO, E. 2009. "Insegnare a tradurre: spunti dall'esperienza in ambito linguistico e apporto delle nuove tecnologie. Il journal di un caso". In *Studi di Glottodidattica. Rivista online del Dipartimento di Pratiche linguistiche e analisi di testi*, 3 (1): 124-156. Università di Bari.
- DI SABATO, B. 2007. "La traduzione e l'apprendimento/insegnamento delle lingue". In *Studi di Glottodidattica*, Vol 1: 47-57.
- DI SPARTI, A. 2004. *Un computer non più calcolatore*. Palermo: Sellerio Editore.

- DI SPARTI, A. 2006. *Tradurre – I cocci di Babele*. Palermo: Edizioni Compostampa.
- DOHERTY, S. e J. Moorkens. 2013. “Investigating the Experience of Translation Technology labs: pedagogical implication”. In *Journal of Specialised Translation*, Num. 19: 122-135
- DOHERTY, S., S. O’Brien; M. Carl. 2010. “Eye tracking as an MT evaluation technique”. In *Machine Translation*, 24 (1): 1-13.
- DRAGSTED, B. 2004. *Segmentation in Translation and Translation Memory Systems: An Empirical Investigation of Cognitive Segmentation and Effects of Integrating a TM System into the Translation Process*. Tesi di dottorato, Copenhagen Business School, Department of International Business Communication.
- DRAGSTED, B. 2006. “Computer-Aided Translation as a Distributed Cognitive Task”. In *Pragmatics & Cognition*, 2006, 14 (2): 443-464.
- ENRIQUEZ RAIDO, V. e F.Austermühl. 2003. “Translation, Localization, and Technology - Current Developments”, In: Pérez-González, Luís (ed.), *Speaking in Tongues: Language Across Contexts and Users.*, University of Valencia Press: 225-250.
- ERICSSON, K. A. e H. A. Simon. 1984. *Protocol Analysis. Verbal Reports as Data*. MIT.
- ERICSSON, K. A. e W. Kintsch. 1995. “Long – term working memory”. In *Psychological Review*, Vol. 102: 211-245.
- ESSELINK, B. 2003. “The evolution of localization.” In *The Guide from Multilingual Computing & Technology: Localization* (57): 21-29.
- ESSELINK, B. 2000. *A Practical Guide to Localization*. Amsterdam/Philadelphia, John Benjamins.
- FERNANDEZ-RODRIGUEZ, M. 2010. “Evolución de la traducción asistida por ordenador. De las herramientas de apoyo a las memorias de traducción”. In *SENDEBAR*, Vol. 21: 201-230
- FOLARON, D. 2010 (a). “Translation Tools”. In *Handbook of Translation Studies*, Vol. 1: 3429-436. Amsterdam/Philadelphia, John Benjamins.
- FOLARON, D. 2010 (b). “Web and Translation”. In *Handbook of Translation Studies*, Vol. 1: 446-450. Amsterdam/Philadelphia, John Benjamins.

- FORCADA, M. 2010. "Machine Translation Today. In *Handbook of Translation Studies*, Vol. 1: 215-223. Amsterdam/Philadelphia, John Benjamins.
- FRASCA, G. 2001. *Videogames of the Oppressed: Videogames as a Means for Critical Thinking and Debate*. Georgia Institute of Technology.
- FRASER, A. e D. Marcu, D. 2007. "Measuring Word Alignment Quality for Statistical Machine Translation". In *Computational Linguistics*, 33, (3):293-303. Disponibile all'indirizzo <http://dl.acm.org/citation.cfm?id=1288683>
- FROELIGER, N. e J. R. Ladmiral. 2010. "De la localisation à la delocalization – le facteur local en traduction". In *Meta*, 55 (4): 615-625. Disponibile all'indirizzo <http://id.erudit.org/iderudit/045681ar>.
- FULFORD, H. e J. Granell Zafra. 2004. "The uptake of Online Tools and Web-based Language Resources by Freelance Translators: Implications for Translator Training, Professional Development and Research". In E. Yuste Rodrigo, (ed.), *Proceedings of the 2<sup>nd</sup> International Workshop on Language Resources for Translation Work, Research and Training (LR4Trans)*, Ginevra: 37-44.
- GALE, W. e K. Church.1993. "A Program for Aligning Sentences in Bilingual Corpora" In *Computational Linguistics*, 19 (1): 75-102. Disponibile all'indirizzo <http://acl.ldc.upenn.edu/J/J93/J93-1004.pdf>
- GANDIN, S. 2009. "Linguistica dei corpora e traduzione: definizioni, criteri di compilazione e implicazioni di ricerca dei corpora paralleli". In *Annali della Facoltà di Lingue e Letterature Straniere dell'Università di Sassari*, Vol. 5: 133-152.
- GASPARI, F. 2001. "Teaching Machine Translation to Trainee Translators: a Survey of Their Knowledge and Opinions". In M. L. Forcada; J. A. Pérez-Ortiz and D. R. Lewis (eds), *Proceedings of the Workshop Teaching Machine Translation. Machine Translation Summit VIII*. Santiago de Compostela: 35-44.
- GOTTI, F.; P. Langlais; D. Bourigault; C. Coulombe. 2005. "3GTM: A Third-Generation Translation Memory" In *3<sup>rd</sup> Computational Linguistics in the North-East*, Gatineau, Québec. Disponibile all'indirizzo <http://www.iro.umontreal.ca/~felipe/Papers/paper-cline-3gtm-2005.pdf>
- GOTTI, M. 2008 *Investigating Specialized Discourse*. Berna, Peter Lang.

- GRAHAM, Y. et al. 2012. "Measurement of Progress in Machine Translation". In *Proceedings of Australasian Language Technology Association Workshop*: 70-78
- HAJIC, J.; P. Homola; V. Kubon. 2003. "A Simple Multilingual Machine Translation System". In *Proceedings of the MT Summit IX*. Disponibile all'indirizzo <http://www.mt-archive.info/MTS-2003-Hajic.pdf>
- HAYNES, C. 1998. *Breaking Down the Language Barriers*. Londra: Aslib
- HEWING, M. 2005. *Advanced Grammar in Use*. Cambridge University Press.
- HODÁSZ, G.; T. Gr bler; B. Kis. 2004. "Translation memory as a robust example-based translation system". In *Proceedings of the Ninth EAMT workshop*, University of Malta, Valletta: 82-89.
- HOLMES, J. 1975. "The Name and Nature of Translation Studies." In James S. Holmes (ed.), *Translated! Papers on Literary Translation and Translation Studies*, Amsterdam: Rodopi: 67–80.
- HUTCHINS, J. 2005. "Current commercial machine translation systems and computer-based translation tools: system types and their uses". In *International Journal of Translation*, 17 (1-2): 5-38
- HUTCHINS, J. 2003. "The Development and Use of Machine Translation Systems and Computer-based Translation Tools". In *International Journal of Translation*, 15 (1): 5-26
- HUTCHINS, J. 2001. "Machine Translation and Human Translation: in Competition or in Complementation?" In *International Journal of Translation*, 13 (1-2): 5-20.
- HUTCHINS, J. 1997. "Evaluation of Machine Translation and Translation Tools". In G.B. Varile, A. Zampolli (eds.) *Survey of the State of the Art in Human Language Technology*: 418-419
- HUTCHINS, J. 1995. "Machine Translation: A Brief History". In E. F. K. Koerner, R. E. Asher (eds.) *Concise History of the Language sciences: from the Sumerians to the Cognitivists*. Oxford, Pergamon Press: 431-445.
- HUTCHINS, J. e H. Somers. 1992. *An Introduction to Machine Translation*. Academic Press, Londra.
- JUOLA, P. 1995. *Learning to Translate: A psycholinguistic approach to the Induction of Grammar and Transfer Functions*. University of Colorado at Boulder.

- KAY, M. 1980. "The Proper Place of Men and Machines in Language Translation". Relazione tecnica CSL – 80 – 11 per Xerox Palo Alto Research Center, California.
- KENNY, D. e A. Way. 2001. "Teaching Machine Translation & Translation Technology: A Contrastive Study". *MT Summit VIII Workshop on Teaching Machine Translation*, Santiago de Compostela: 13–17.
- KENNY, D. 1999. "CAT Tools in an Academic Environment: What Are They Good For?" In *Target*, 11 (1): 65-82.
- KILGRAY Translation Technologies. 2004 – 2013. *MemoQ Quick Start Guide*.
- KIRALY, D. C. 1995. *Pathways to Translation. Pedagogy and Process*. Kent State University Press.
- KOEHN, P. 2005. "Europarl: A Parallel Corpus for Statistical Machine Translation". In *MT Summit*, vol. 5: 79-86.
- KOEHN, P. et al. 2007. "Moses: Open source toolkit for statistical machine translation". In *Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions*. Association for Computational Linguistics: 177-180.
- KOEHN, P. 2004. "Statistical Significance Tests for Machine Translation Evaluation". In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP)*: 388–395.
- KOSMA, A. 2007. "Le fonctionnement spécifique de la mémoire de travail en traduction". In *Meta*, 52 (1): 22-28. Disponible all'indirizzo <http://id.erudit.org/iderudit/014716ar>.
- KRINGS, H. 2001. *Repairing Texts: Empirical Investigations of Machine Translation Post-Editing Processes*. G. S. Koby (ed.), Kent State University.
- KROLLMANN, F. 1981. "Computer Aids to Translation". In *Meta*, 26 (1): 85-94. Disponibile all'indirizzo <http://id.erudit.org/iderudit/003353ar>.
- LADMIRAL, J. R. 2010. "La traduction, phénomène interculturel et psychorelationnel". In *Meta*, 55 (4): 626-641. Disponibile all'indirizzo <http://id.erudit.org/iderudit/045682ar>
- LADMIRAL, J. R. 2004. "Dichotomies Traductologiques". In *La linguistique* (40): 25-50

- LAGOUDAKI, E. 2007. “Challenges and Possibilities for Extracting Parallel Corpora from the Web – The Translator’s Dream Scenario”. In *X Simposio Internacional Comunicación Social*, Santiago de Cuba: 474-479.
- LAGOUDAKI, E. 2006. “Translation Memories Survey 2006: Users’ Perception around TM Use”. In *Proceedings of the International Conference Translating and the Computer 28*. London, Aslib.
- LAROCHE, A. e P. Langlais. 2010. “Revisiting context-based projection methods for term-translation spotting in comparable corpora”. In *Proceedings of the 23rd international conference on computational linguistics*. Association for Computational Linguistics: 617-625.
- LEVITT, G. 2003. “Internet-Based Sharing of Translation Memory”. In *MultiLingual Computing & Technology*, 14 (5): 38-41
- LI, D. 2002. “Translator Training: What Translation Students Have to Say”. In *Meta*, 47 (4): 513-531. Disponibile all’indirizzo <http://id.erudit.org/iderudit/008034ar>.
- LIDDY, E. D. 2001. “Natural Language Processing”. In *Encyclopedia of Library and Information Science*, 2° ed., Marcel Decker, Inc., New York.
- LOPEZ, A. 2008. “Statistical Machine Translation”. In *ACM Computing Surveys*, 40 (3), Articolo 8.
- MACKLOVITCH, E.; M. Simard; P. Langlais. 2000. “TransSearch: A Free Translation Memory On The World Wide Web”, In *2<sup>nd</sup> International Conference on Language Resources and Evaluation (LREC’00)*, Atene; 1201–1208.
- MAIA, B. 2005. “Terminology and Translation — bringing research and professional training together through technology”. In *Meta*, 50 (4). Disponibile all’indirizzo <http://id.erudit.org/iderudit/019921ar>
- MANDELBLIT, N. 1993. “Machine Translation: a Cognitive Linguistics Approach”. In *Proceedings of the 5th Int. Conf. on Theoretical and Methodological Issues in Machine Translation*, Kyoto: 117-130.
- MARCU, D. e W. Wong. 2002. “A Phrase-Based, Joint Probability Model for Statistical Machine Translation”. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Philadelphia: 133-139.

- MARTÍN, R. M. 2007. “Traductología cognitiva y traductología empírica”. In G. Wotjac (ed.) *Quo vadis Translatologie?* Berlino, Frank & Timme
- MAZZOTTA, P. 2007. “Riflessioni glottodidattiche sulla traduzione delle microlingue”. In Montella, Marchesini (a cura di), *I saperi del tradurre*. Milano, Franco Angeli
- MCDONOUGH, J. 2007. “How Do Language Professionals Organize Themselves? An Overview of Translation Networks”. In *Meta* 52 (4): 793-815. Disponibile all’indirizzo <http://id.erudit.org/iderudit/017697ar>.
- MELBY, A. K. 2006. “MT+TM+QA: The Future is Ours” In *Tradumàtica: traducció i tecnologies de la informació i la comunicació* (4).
- MELBY, A. K. 1981. “Translators and Machines – Can they Cooperate?” In *Meta* 26 (1): 23-34. Disponibile all’indirizzo <http://id.erudit.org/iderudit/003619ar>.
- MITAMURA, T. e E. Nyberg. 1995. “Controlled English for Knowledge-Based MT: Experience with the KANT System”. In *Proceedings of the Sixth International Conference on Theoretical and Methodological Issues in Machine Translation*. 5-7 Luglio 1995, Leuven, Belgio: 158-172.
- MITAMURA, T.; E. Nyberg; J. Carbonell. 1991. “An Efficient Interlingua Translation System for Multi-lingual Document Production”. In *Proceedings of Machine Translation Summit III*, Washington D.C.
- MOORKENS, J. 2012. *Measuring Consistency in Translation Memories: a Mixed-Methods Case Study* (Tesi di dottorato, Dublin City University).
- MOSER-MERCER, B. 1987. “Man/Machine Interface in Translation and Terminology”. In *Meta* 32 (2): 156-163. Disponibile all’indirizzo <http://id.erudit.org/iderudit/003619ar>.
- MOSSOP, B. 2007. “Empirical Studies of Revision: What We Know and Need to Know”. In *The Journal of Specialised Translation*, 8: 5-20.
- MOSSOP, B. 2006. “Has Computerization Changed Translation?”. In *Meta* 51 (4): 787-805. Disponibile all’indirizzo <http://id.erudit.org/iderudit/014342ar>.
- MUNDAY, J. (ed.). 2009. *The Routledge Companion to Translation Studies*. Londra / New York: Routledge.
- NAGATA, M., T. Saito e K. Suzuki. 2001. “Using the web as a bilingual dictionary”. In *Proceedings of the workshop on Data-driven methods in machine translation* (14) Association for Computational Linguistics: 1-8.

- NEWMARK, P. 1988. *A Textbook of Translation*. Londra, Prentice Hall
- NIRENBURG, S. 1996. “Bar Hillel and Machine Translation: Then and Now”. In *Koppel and Shamir*: 300-305.
- NIRENBURG, S. 1992. “Tools for Machine-Aided Translation: The CMU TWS”. In *Meta* 37 (4): 709-720. Disponibile all’indirizzo <http://id.erudit.org/iderudit/003739ar>.
- O’BRIEN, S. (ed.). 2011. *Cognitive Explorations of Translation*. London and New York: Continuum.
- O’BRIEN, S., M. O’Hagan, e M. Flanagan. 2010. “Keeping an Eye on the UI Design of Translation Memory: How do Translators Use the Concordance Feature?” In *ACM, Proceedings of the 28th Annual European Conference on Cognitive Ergonomics*: 187-190.
- O’BRIEN, S. 2009. “Eye tracking in translation process research: methodological challenges and solutions”. In *Methodology, Technology and Innovation in Translation Process Research*, 38: 251-266.
- O’BRIEN, S. 2008. “Processing fuzzy matches in translation memory tools: An eye-tracking analysis”. In *Looking at eyes—eye tracking studies of reading and translation processing*. Copenhagen Studies in Language, 36: 79-102.
- O’BRIEN, S. 2006. “Pauses as Indicators Of Cognitive Effort In Post-Editing Machine Translation Output”. In *Across Languages and Cultures* 7 (1). Akadémiai Kiadó, Budapest: 1 - 21
- O’HAGAN, M. 2006. “Teletranslation Revised: Futurama for Screen Translators?” In *MuTra – Audiovisual Translation Scenarios: Conference Proceedings*: 1-12
- OCH, F. J. e H. Ney. 2004. “The Alignment Template Approach to Statistical Machine Translation”. In *Computational Linguistics* 30 (4): 417-449
- OLOHAN, M. 2011. “Translators and Translation Technology: The Dance of Agency”. In *Translation Studies* 4 (3): 342-357.
- OLVERA, M. et al. (2005): “Translator training and modern market demands”. In *Perspectives*, 13 (2): 132-142
- OROZCO, M. e A. Hurtado Albir. 2002. “Measuring Translation Competence Acquisition”. In *Meta*, 47 (3): 375-402. Disponibile all’indirizzo <http://id.erudit.org/iderudit/008022ar>.

- PACTE. 2008. "First Results of a Translation Competence Experiment: 'Knowledge of Translation' and 'Efficiency of the Translation Process'". In J. Kears (ed.), *Translator and Interpreter Training*. Londra / New York, Continuum: 104-126.
- PACTE. 2007. "Une recherche empirique expérimentale sur la compétence en traduction". In D. Gouadec (ed.) *Actes du Colloque International : Quelle formation pour le traducteur?*. Parigi, Maison du dictionnaire: 95-116.
- PACTE. 2005. "Investigating Translation Competence: Conceptual and Methodological Issues". In *Meta*, 50 (2): 609-619.
- PACTE. 2003. "Building a Translation competence Model". In F. Álvarez (ed.) *Triangulating Translation: perspectives in process oriented research*, Amsterdam/Philadelphia, John Benjamins: 43-68.
- PALUMBO, G. 2009. *Key Terms in Translation Studies*. Londra/New York, Continuum.
- PAPAGNO, C. 2003. *Come funziona la memoria*. Bari: Editori Laterza.
- PAPINENI, K., S. Roukos, T. Ward e W. J. Zhu. 2002. "Bleu: A Method for Automatic Evaluation of Machine Translation". In *Proceedings of the 40<sup>th</sup> Annual Meeting of the Association of Computational Linguistics*, Philadelphia: 311 -318
- PARDELLI, G. e S. Biagioni. 2013. "Quando la linguistica incontra l'informatica: una riflessione terminologica". In *SCIRES-IT*, 3 (1): 67-78. CAPSUR-CIBER Publishing.
- PIERCE, J. R., et al. 1966. *Computers in translation and linguistics* (ALPAC report). National Academy of Sciences/National Research Council.
- PLANAS, E. e O. Furuse. 1999. "Formalizing Translation Memories". In *MT Summit VII*: 331-339
- POPOVIC, M. e N. Hermann. 2011. "Towards Automatic Error Analysis of Machine Translation Output". In *Computational Linguistics* 37 (4): 657-688.
- PYM, A. 2013. "Translation Skill-sets in a Machine-translation Age". In *Meta* 58 (3), 487-503.
- PYM, A. 2011 (a). "What Technology Does to Translating". In *The International Journal of Translation and Interpreting*, 3 (1): 1-9.
- PYM, A. 2011 (b). "Democratizing Translation Technologies – The Role of Humanistic Research". Saggio presentato alla Conferenza della Luspio sulla traduzione automatica,

Roma 5 aprile 2001. Disponibile all'indirizzo [http://usuaris.tinet.cat/apym/online/research\\_methods/2011\\_rome.pdf](http://usuaris.tinet.cat/apym/online/research_methods/2011_rome.pdf).

PYM, A. 2010. *Exploring Translation Theories*. Londra/New York, Routledge.

PYM, A. 2006. "Globalization and the Politics of Translation Studies". In *Meta* 51 (4): 744-757. Disponibile all'indirizzo <http://id.erudit.org/iderudit/014339ar>.

PYM, A. 2003. "Redefining Translation Competence in an Electronic Age. In Defence of a Minimalist Approach". In *Meta* 48 (4): 481-497. Disponibile all'indirizzo <http://id.erudit.org/iderudit/008533ar>.

QUAH, C. K.. 2006. *Translation and Technology*. Hampshire/New York, Palgrave Mcmillan.

QUIRION, J. 2003. "La formation en localisation à l'université: pour quoi faire?" In *Meta* 48 (4): 546-558. Disponibile all'indirizzo <http://id.erudit.org/iderudit/003619ar>.

RABADAN, R. 2010. "Applied Translation Studies". In *Handbook of Translation Studies*, Vol. 1: 7-11. Amsterdam/Philadelphia, John Benjamins.

RAWSON, K. e J. Van Overschelde. 2008. "How Does Knowledge Promote Memory? The Distinctiveness Theory of Skilled Memory". In *Elsevier – Journal of Memory and Language*. Vol. 58: 646-668.

REINKE, U. 2013. "State of the Art in Translation Memory Technology". In G. Rehm, F. Sasaki, D. Stein, A. Witt (eds.), *Translation: Computation, Corpora, Cognition. Special Issue on Language Technologies for a Multilingual Europe*, 3 (1): 27-48.

REPOVS, G. e BADDELEY, A. D. 2006. "The Multi-Component Model of Working Memory: Explorations in Experimental Cognitive Psychology". In *Neuroscience*, 139: 5 – 21.

REX RICKS, T. e J. Wiley. 2009. "The Influence of Domain Knowledge on the Functional Capacity of Working Memory". In *Elsevier – Journal of Memory and Language*. Vol. 61: 519-537.

RICO PEREZ, C. e A. Martín de Santa Olalla Sánchez. 1997. "New Trends in Machine Translation" In *Meta* 42 (4): 605-615. Disponibile all'indirizzo <http://id.erudit.org/iderudit/003822ar>.

RIESA, J; A. Irvine; D. Marcu. 2011. "Feature-Reach Language-Independent Syntax-Based Alignment for Statistical Machine Translation". In *Proceedings of the 2011*

*Conference on Empirical Methods in Natural Language Processing*. Association of Computational Linguistics, Edimburgo, Scozia (UK): 497-507.

ROBERT, I. 2008. "Translation Revision Procedures: An Explorative Study". In *Translation and Its Others. Selected Papers of the CETRA Research Seminar on Translation Studies*.

RONOWICZ, E; J. Hehir; T. Kaimi; K. Kojima; D. Lee. 2005. "Translator's Frequent Lexis Store and Dictionary Use as Factors in SLT Comprehension and Translation Speed – A Comparative Study of Professional, Paraprofessional and Novice Translators". In *Meta* 50 (2): 580-596.

SAGER, J. C. 1994. *Language Engineering and Translation: Consequences of Automation*. Amsterdam/Philadelphia, John Benjamins.

SALMON, L. 2005. "Proposta teorica sui processi traduttivi umani", in G. Garzone (a cura di), *Esperienze del tradurre. Aspetti teorici e applicativi*, Milano, Franco Angeli: 15-34

SALMON, L. 2007. "Sulla gestione del processo traduttivo umano: cosa sappiamo fare senza sapere come". In C. Montella e G. Marchesini (a cura di), *I saperi del tradurre. Analogie, affinità, confronti*. Milano, Franco Angeli: 71-100.

SALMON, L. e M. Mariani. 2008. *Bilinguismo e traduzione. Dalla neurolinguistica alla didattica delle lingue*. Milano, Franco Angeli.

SAMPSON, G. e D. McCarthy (ed). 2004. *Corpus Linguistics: Readings in a Widening Discipline*. Londra e New York: Continuum.

SASSI, M. e M. L. Ceccotti. 2000. "L'utilizzo didattico di corpora: proposte metodologiche". In *Proceeding of Didamatica 2001, Informatica per la Didattica*. Disponibile all'indirizzo <http://www.ilc.cnr.it/CEG/progetto/pdf/Linguistica%20dei%20corpora.pdf>

SCHALER, R. 2010. "Localization and Translation Studies". In *Handbook of Translation Studies*, Vol. 1: 209-214. Amsterdam/Philadelphia, John Benjamins.

SCHALER, R. 2001. "Beyond Translation Memories". In *Workshop on Example-Based Machine Translation*: 49-55.

- SCHERF, W. G. 1992. "Training, Talent and Technology." In C. Dolleroup e A. Loddegaard (ed.), *Teaching Translation and Interpreting: Training, Talent and Experience*: 153-160. Amsterdam/Philadelphia, John Benjamins.
- SCHILPEROORD, J. 1996. *It's About Time: Temporal Aspects of Cognitive Processes in Text Production*. Amsterdam: Rodopi.
- SCHNEIDER, T. 1983. "Some Notes on Machine Aids for Translators". In *Meta* 28 (4): 344-351. Disponibile all'indirizzo <http://id.erudit.org/iderudit/004048ar>.
- SDL TRADOS. 2011. Multiterm Extract User Guide.
- SDL TRADOS. 2007. Translator's Workbench – Getting Started Manual.
- SDL TRADOS. 2007. Translator's Workbench – Intermediate Manual.
- SDL TRADOS. 2014. SDL Trados Studio 2014 – Getting Started Manual.
- SDL TRADOS. 2014. SDL Multiterm Desktop 2014 –Manual for Translators.
- SEGAL, A. 1991. "Machine Translation of Natural Languages: the TOVNA MT Solution: A Learning System". In *Translating and the Computer, 12: Applying Technology to the Translation Process*. Londra, Aslib: 83-109.
- SERETAN, V. 2010. "An integrated environment for extracting and translating collocations". In M. Mahlberg, V. Gonzalez-Diaz e C. Smith (eds), *Proceedings of the Corpus Linguistics Conference, CL2009*, Liverpool. Disponibile all'indirizzo <http://www.latl.unige.ch/personal/vseretan/>
- SERETAN, V. e E. Wehrli. 2010. "Tools for Syntactic Concordancing". In *Proceedings of the International Multiconference on Computer Science and Information Technology*, Vol. 5: 493–500
- SERETAN, V. 2009. "Extraction de collocations et leurs équivalents de traduction à partir de corpus parallèles". In *TAL*, 50 (1): 305–332. Disponibile all'indirizzo <http://www.issco.unige.ch/en/staff/seretan/publ/TAL-2009-50-1-12-Seretan.pdf>
- SERETAN, V. e E. Wehrli. 2007. "Multilingual Collocation Extraction with a Syntactic Parser". In *Language Resources and Evaluation*, 43 (1): 71–85. Disponibile all'indirizzo <http://www.latl.unige.ch/personal/vseretan/>
- SHAROFF, S. 2004. "Translation as problem solving: uses of comparable corpora". In Yuste Rodrigo, Elia (ed), *Proceedings of the Third International Workshop on Language*

*Resources for Translation Work, Research & Training (LR4Trans-III)*, Parigi, ELRA (European Language Resources Association).

SHAROFF, S.; B. Babych; A. Hartley. 2007. “Irrefragable answers using comparable corpora to retrieve translation equivalents”. In *Language Resources and Evaluation* 43 (1): 15-25. Disponibile all’indirizzo <http://www.springerlink.com/content/8k6631431pl35381/>.

SHREVE, G. 2006. “The Deliberate Practice: Translation and Expertise”. In *Journal of Translation Studies* 9 (1): 27-42.

SIMARD, M. e P. Isabelle. 2009. “Phrase-Based Machine Translation in a Computer-Assisted Translation Environment”. In *Proceedings of the Twelfth Machine Translation Summit*: 120-127.

SOMERS, H. (ed). 2003. *Computers and Translation: A Handbook for Translators*. Amsterdam/Philadelphia, John Benjamins.

SPECIA, L. *et al.* 2009. “Estimating the Sentence-Level Quality of Machine Translation Systems”. In *Proceedings of the 13<sup>th</sup> Annual Conference of the EAMT*: 28–35.

STEIN, D. 2013. “Machine Translation. Past, Present and Future”. In G. Rehm, F. Sasaki, D. Stein, A. Witt (eds.), *Translation: Computation, Corpora, Cognition. Special Issue on Language Technologies for a Multilingual Europe*, 3 (1): 5-13.

STOLL, C.H. 1988. “Translation Tools on PC”. In C. Picken (ed.) *Translating and the Computer 9: Potential and Practice*. Londra, Aslib: 11-26

TAN, Z. 2008. “Towards a Whole-Person Translator Education Approach in Translation Teaching on University Degree Programmes”. In *Meta* 53 (3): 589-608. Disponibile all’indirizzo <http://id.erudit.org/iderudit/019241ar>.

TARAVELLA, A.M. e A. O. Villeneuve. 2013. “Acknowledging the Needs of Computer-Assisted Translation Tools Users: The Human Perspective in Human-Machine Translation”. In *Journal of Specialized Translation*, Num. 19: 62-74

THICKE, L. 2013. “The Industrial Process for Quality Machine Translation” In *Journal of Specialised Translation*, Vol. 19: 8-17

THURMAIR, G. 2005. “Hybrid Architectures for Machine Translation Systems”. In *Language Resources and Evaluation*, 39 (1): 91-108 Disponibile all’indirizzo <http://link.springer.com/article/10.1007/s10579-005-2698-z#page-1>

- TUFIS, D.; A. M. Barbu; R. Ion. 2004. “Extracting Multilingual Lexicons from Parallel Corpora”. In *Computers and Humanities* 38: 163–189. Kluwer Academic Publishers, Netherlands.
- ULRYCH, M. (ed.) 1997. *Tradurre. Un approccio multidisciplinare*. UTET Torino.
- VANDAELE, S. e L. Lubin. 2005. “Approche cognitive de la traduction dans les langues de spécialité: vers une systématisation de la description de la conceptualisation métaphorique”. In *Meta* 50 (2): 415-431. Disponibile all’indirizzo <http://id.erudit.org/iderudit/010991ar>
- VULI, I. e M. F. Moens. 2012. “Detecting highly confident word translations from comparable corpora without any prior knowledge”. In *Proceedings of the 13<sup>th</sup> Conference of the European Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics: 449-459.
- WALLIS, J. 2008. “Interactive Translation vs. Pre-Translation in TMs: A Pilot Study”. *Meta* 53 (3): 623-629. Disponibile all’indirizzo <http://id.erudit.org/iderudit/019243ar>.
- WALLIS, J. 2006. *Interactive Translation vs. Pre-Translation in the Context of Translation Memory Systems: Investigating the Effects of Translation Method on Productivity, Quality and Translators Satisfaction*. (Tesi di dottorato, School of Translation and Interpretation, University of Ottawa).
- WANG, W.; J. May; K. Knight; D. Marcu. 2010. “Re-structuring, Re-labeling and Re-aligning for Syntax-Based Machine Translation”. In *Computational Linguistics*, 36 (2): 247-277.
- WEAVER, W. 1955. “Translation” (1949). In *Machine Translation of Languages*. Cambridge MA, MIT Press.
- WHERLI, E.; V. Seretan; L. Nerima; L. Russo. 2009. “Collocations in a Rule-Based MT System: A Case Study Evaluation of Their Translation Adequacy”. In *Proceedings of the 13th Annual Conference of EAMT (European Association for Machine Translation)*: 128-135. Disponibile all’indirizzo [http://www.latl.unige.ch/personal/vseretan/publ/EAMT2009\\_EW\\_VS\\_LN\\_LR.pdf](http://www.latl.unige.ch/personal/vseretan/publ/EAMT2009_EW_VS_LN_LR.pdf)
- WILKS, Y. 1972. *An Artificial Intelligence Approach to Machine Translation*. Stanford University, Computer Science Department.

- WINKLER, R. P.; S. Metu; S. La Rocca; J. Morgan. 2013. *Adding Statistical Machine Translation Adaptation to Computer-Assisted Translation*. Army Research Laboratory, Adelphi MD Computational and Information Sciences Directorate.
- WOTJAK, G. 1997 "Problem-Solving Strategies in Translation". In *Revista Ilha do Desterro. A Journal of English Language, Literatures in English and Cultural Studies*, Florianópolis, Universidade Federal de Santa Catarina, Num. 33: 101-116.
- ZANETTIN, F. 2002. "Corpora in translation practice". In E. Yuste-Rodrigo (ed.), *Language Resources for Translation Work and Research, LREC 2002 Workshop Proceedings*, Las Palmas de Gran Canaria. Disponibile all'indirizzo <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.136.5669&rep=rep1&type=pdf>.
- ZIMMERMAN, C. e C. Kelley. 2010. "'I'll remember this!' Effects of emotionality on memory predictions versus memory performance". In *Elsevier - Journal of Memory and Language*, Vol. 62: 240-253.

