



UNIVERSITÀ DEGLI STUDI INTERNAZIONALI DI ROMA (UNINT)

FACOLTÀ DI INTERPRETARIATO E TRADUZIONE

**CORSO DI LAUREA MAGISTRALE IN
INTERPRETARIATO E TRADUZIONE**

TESI DI LAUREA IN TRADUZIONE AUTOMATICA E POST-EDITING

**TRADUZIONE AUTOMATICA E POST-EDITING: ANALISI
COMPARATIVA DELLE PRESTAZIONI DEI NUOVI
CAT TOOL WEB-BASED MATECAT, SMARTCAT E SYSTRANET**

Candidato
Jacopo Santonico

Relatrice
Fiorenza Mileto
Correlatore
Marco Stefano Tomatis

Anno Accademico 2014/15

Sessione straordinaria

INDICE

<i>Abstract</i>	4
<i>Resumen</i>	4
<i>Introduzione</i>	6
1. LA STATISTICAL MACHINE TRANSLATION: LA STORIA, LA CREAZIONE, LA VALUTAZIONE	8
1.1 NASCITA ED EVOLUZIONE DELLA MACHINE TRANSLATION: DAL MEMORANDUM DI WILLIAM WEAVER ALLA STATISTICAL MACHINE TRANSLATION (SMT)	8
1.2 LA STATISTICAL MACHINE TRANSLATION: COS'È E COME SI CREA UN SISTEMA DI TRADUZIONE AUTOMATICA SMT	11
1.2.1 CREARE UN SISTEMA DI TRADUZIONE SMT I: MODELLI DI EQUIVALENZA TRADUTTIVA	13
1.2.2 CREARE UN SISTEMA DI TRADUZIONE SMT II: PARAMETRIZZAZIONE	15
1.2.3 CREARE UN SISTEMA DI TRADUZIONE SMT III: STIMA DEI PARAMETRI.....	18
1.2.4 CREARE UN SISTEMA DI TRADUZIONE SMT IV: DECODING.....	20
1.3 VALUTARE L'OUTPUT DI UN SISTEMA DI TRADUZIONE AUTOMATICA	24
1.3.1 BLEU: BILINGUAL EVALUATION UNDERSTUDY	25
1.3.2 NIST: NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY	26
1.3.3 METEOR: METRIC FOR EVALUATION OF TRANSLATION WITH EXPLICIT ORDERING	26
1.3.4 TER: TRANSLATION EDIT RATE	27
2. IL POSTEDITING: ANALISI DEGLI STANDARD PRODUTTIVI E QUALITATIVI	28
2.1 VALUTARE GLI STANDARD QUALITATIVI DELLA TRADUZIONI AUTOMATICA: PARAMETRI DI MACRO E MICROVALUTAZIONE	28
2.2 POST-EDITING E INCREMENTO DEGLI STANDARD DI PRODUZIONE: METODOLOGIE DI STUDIO DELLA CORRELAZIONE TRA QUALITÀ E PRODUTTIVITÀ	40
2.2.1 POST-EDITING E INDICATORI DI INCREMENTO PRODUTTIVO: ANALISI DELLE PRESTAZIONI IN FUNZIONE DELL'INPUT SU MOUSE E TASTIERA	43
2.2.2 CORRELARE QUALITÀ E SFORZO COGNITIVO NEL POST-EDITING: EYETRACKING E FISSAZIONI OCULARI	47
2.3 POTENZIAMENTO E OTTIMIZZAZIONE DELLE PRESTAZIONI: PRE-EDITING, POST-EDITING AUTOMATIZZATO E CONFIDENCE SCORE	50

2.3.1 PRE-EDITING E OTTIMIZZAZIONE DEI TESTI ATTRAVERSO L'IMPLEMENTAZIONE DI SET DI REGOLE CONTROLLED LANGUAGE (CL)	51
2.3.2 TECNICHE DI AUTOMATIZZAZIONE DEL POST-EDITING E <i>CONFIDENCE SCORES</i> ..	55
2.4 IL POST-EDITOR: PROFILO E REQUISITI.....	59
3. TRADUZIONE AUTOMATICA E POST-EDITING ALL'OPERA: ANALISI PRODUTTIVO-QUALITATIVA DELLE PRESTAZIONI DI MATECAT, SMARTCAT E SYSTRANET	62
3.1 TRADUZIONE AUTOMATICA E TOOL OPEN-SOURCE: ALCUNE CONSIDERAZIONI PRELIMINARI.....	62
3.1.1 CARATTERISTICHE E FUNZIONALITÀ DI MATECAT	65
3.1.2 CARATTERISTICHE E FUNZIONALITÀ DI SMARTCAT	69
3.1.3 CARATTERISTICHE E FUNZIONALITÀ DI SYSTRANET.....	71
3.2. LA TIPOLOGIA TESTUALE ESAMINATA.....	73
3.2.1 L'ITALIANO COMUNITARIO: L'EUROLETTO	75
3.3 COME VALUTARE LA QUALITÀ DELLA MACHINE TRANSLATION: UNA METODOLOGIA DI LAVORO	77
3.3.1 LA PROSPETTIVA MACROVALUTATIVA: ANALISI DEI PARAMETRI DI CHIAREZZA, ACCURATEZZA E STILE.....	80
3.3.2 IMPLEMENTAZIONE DEI PROCESSI DI REVISIONE E QUALITY ASSURANCE: XBENCH E DQF.....	90
3.3.3 TRADUZIONE AUTOMATICA E REFERENCE TRANSLATION A CONFRONTO.....	102
3.4 COME VALUTARE LA PRODUTTIVITÀ DELLA MACHINE TRANSLATION: KEYTRACKER E PRODUCTIVITY EVALUATION (PE)	104
3.4.1 MONITORAGGIO DEL PROCESSO DI POST-EDITING ATTRAVERSO L'UTILIZZO DEL KEYTRACKER INPUTLOG.....	105
3.4.2 DYNAMIC QUALITY FRAMEWORK E MATECAT: CALCOLO DELLA PRODUTTIVITÀ E DEL POST-EDITING EFFORT (PEE).....	111
3.5 DISCUSSIONE DEI RISULTATI.....	114
<i>Conclusioni</i>.....	117
<i>Bibliografia</i>	120
<i>Appendici</i>.....	127
APPENDICE A	128

APPENDICE B.....	134
APPENDICE C	140
APPENDICE D	146
APPENDICE E.....	152

Abstract

Today, machine translation and post-editing are synonyms of fast turn-around and productivity increase, to the detriment of translation quality. In spite of the fact that, most time, machine translation fails to meet quality criteria, iMT technologies prove to be an asset for the translator rather than a hindrance. This assertion is now widely acknowledged in the industry and it reflects itself with the proliferation of new CAT tools and innovations in the translation marketplace. As a consequence, therefore, the market has shown a growing interest for lower-cost, faster turn-around translation services, whose quality standards, however, do not seem to keep up with the pace. In order to monitor and test the performance of these web-based, open-source, newly developed tools and assess the effectiveness of this increasing trend, the setting up of a quality evaluation methodology seems of pivotal importance. This task has been fulfilled analysing the performances of three free, web-based CAT tools, carrying out quality evaluation tasks in terms of Clarity, Accuracy, Adequacy, Terminology and Style, as well as productivity degree and post-editing effort. Productivity measurements have been carried out through a keystroke tracker. One tool out of the three tested has shown that quality is of acceptable level with consequent productivity gain and large margins of improvement, while the third tool tested did not prove to be generally up to the task. These results acknowledge the high potential of these new resources as well as the innovative impulse that these new technologies can bring to the research and the market.

Resumen

Hoy, la traducción automática y el post edición son percibidos como sinónimos de tiempos de entregas reducidos y incremento de la productividad, aunque se experimente una consiguiente pérdida de calidad de la traducción. A pesar de que, la mayoría de las veces, la traducción automática no cumple con los criterios de calidad, más que un obstáculo, la tecnología MT demuestra su capacidad de ser una ayuda para el traductor. Esta afirmación ahora es ampliamente reconocida en la industria y los resultados se reflejan en la proliferación en el mercado de una multitud de programas para la traducción asistida por computadora (TAC) y automática, juntos a diferentes innovaciones tecnológicas. Por consiguiente, el

mercado ha mostrado un interés creciente para programas y servicios de traducción de bajo costo y tiempos de entrega reducidos, cuya calidad, sin embargo, todavía no se ajusta a la tendencia y al ritmo de desarrollo actual de la industria. Para monitorizar y evaluar el rendimiento de estos recién creados programas gratuitos y basados en web y valorar la efectividad de esta nueva tendencia, desarrollar una metodología de valuación de la calidad parece una tarea de importancia fundamental. Esta tarea ha sido cumplida analizando el rendimiento de tres programas gratuitos basados en web llevando a cabo una valuación de la calidad en términos de Claridad, Precisión, Adecuación, Terminología y Estilo, además de nivel de productividad y esfuerzo en post edición. El nivel de productividad ha sido medido a través de un programa de monitorización y registración de las pulsaciones dada en el teclado. Entre los tres, un programa ha demostrado una calidad de nivel suficiente y, por lo tanto, una consiguiente gana en términos de productividad, juntos a grandes márgenes de mejora, mientras que el tercero programa no logró cumplir con la tarea asignada. Estos resultado reconocen el alto potencial de esto nuevos medios, además de los impulsos innovadores que esta tecnología podría llevar a la investigación y al mercado.

Introduzione

Lo scorso ottobre a San Jose, California, si è tenuta l'annuale conferenza TAUS avente per temi l'industria della traduzione, l'attuale panorama in cui si trova e le innovazioni che di volta in volta si riversano sul mercato. Un importante focus è stato ovviamente posto sull'attuale era di grandi innovazioni tecnologiche e su interazioni ed effetti che questa ha sull'industria della traduzione. Come è ormai evidente a tutti, nel mercato globalizzato la traduzione diventa ogni giorno una necessità sempre maggiore, perfettamente integrata, grazie agli indiscutibili progredimenti tecnologici compiuti, in ogni dispositivo, schermo o applicazione. Ciononostante, oggi la mole di contenuti non ancora tradotti rimane considerevole, motivo per cui diviene di fondamentale importanza approntare nuove metodologie e strategie di lavoro. La traduzione automatica sembra la risposta migliore a questa esigenza. Nonostante un atteggiamento in parte ancora restio all'automatizzazione, tanto nel settore della traduzione quanto in altri, e limiti oggettivi tuttora presenti, in virtù delle enormi potenzialità, l'industria della traduzione è oggi pervasa, come mai prima d'ora, da grandi ottimismo ed entusiasmo (Marcu, 2015). Partendo dal presupposto che, almeno per ora, l'eventualità che la traduzione automatica si sostituisca interamente al traduttore rimane pura fantascienza, l'implementazione dei sistemi MT oggi avviene tramite il crescente supporto che questa tecnologia offre nel flusso di lavoro quotidiano di un traduttore, consentendo un considerevole aumento della produttività.

Tra i trend odierni attualmente in continua crescita, uno di questi vede il proliferare sul mercato di strumenti che offrono servizi a costi moderati e con tempi di consegna notevolmente ridotti. È questo quindi il caso di applicazioni disponibili sul mercato gratuitamente o in modalità open-source, tramite cui garantire accesso a numeri imponenti di utenti che si servono di suddetti tool per completare i propri contratti di lavoro. È questo ad esempio il caso di Matecat, sviluppato da Translated.net (ivi). Di pari passo a questi sistemi, ve ne sono altri che invece puntano alla creazione di nuovi modelli lavorativi che sfruttano a pieno le possibilità del nuovo mercato del lavoro in rete, sviluppando piattaforme di lavoro pensate per l'outsourcing e la creazione di reti di traduttori freelance e Language Service Provider impegnati, in maniera collaborativa, nella gestione e lavorazione di progetti

sullo stesso tool gratuito e web-based. È questo il caso di SmartCAT, progettato da Abbyy (ivi).

Del resto, con un interesse la cui portata è proporzionale al crescente utilizzo della traduzione automatica, lo sguardo attento del settore è riservato anche alla delicata questione della qualità traduttiva. Posto che gli attuali sistemi MT non sono ancora in grado di restituire standard di qualità elevati o quantomeno costanti, è diventato di fondamentale importanza approntare metodologie di valutazione che permettano di stimare la qualità dell'output della machine translation, al fine di individuare e distinguere, dal punto di vista qualitativo, un prodotto tradotto automaticamente di buona qualità da uno di qualità insufficiente, onde evitare che testi di scarsa qualità si riversino nel mercato a danno dei post-editor e della loro produttività. In quest'ottica, sono diverse le società che recentemente commercializzano servizi e tool che consentono a traduttori freelance e post-editor di stimare la qualità del testo tradotto automaticamente e il conseguente livello di produttività, nonché i margini di guadagno (ivi).

A tal fine, in questa sede si cercherà di offrire una panoramica quantomeno esaustiva di metodologie di valutazione della qualità e della produttività della machine translation, applicando queste alle prestazioni di tre tool disponibili gratuitamente sul mercato, selezionati in base a funzionalità e modalità di fruizione e utilizzo simili tra loro. Mentre il primo capitolo offrirà un breve excursus su genesi e storia della traduzione automatica e delle *automatic metric* di valutazione, nel secondo capitolo si esploreranno proposte di tassonomie e metodologie di valutazione della MT e dell'attività di post-editing. Il terzo capitolo conterà della parte più squisitamente applicativa di questo studio, poiché vedrà l'applicazione dei processi di *Quality assurance* all'output di tre sistemi MT diversi nella traduzione di testi giuridici comunitari, i quali andranno a costituire l'appendice conclusiva a questo studio.

1. LA STATISTICAL MACHINE TRANSLATION: LA STORIA, LA CREAZIONE, LA VALUTAZIONE

1.1 NASCITA ED EVOLUZIONE DELLA MACHINE TRANSLATION: DAL MEMORANDUM DI WILLIAM WEAVER ALLA STATISTICAL MACHINE TRANSLATION (SMT)

Nel tentativo di rintracciare l'esatto momento della storia in cui la traduzione automatica ha avuto inizio, molti studiosi convergono nell'identificare il ventesimo secolo come punto d'origine, alla luce dei primi effettivi esperimenti pratici, sebbene quella di rudimentali dizionari meccanici fosse un'idea che iniziò a circolare addirittura nel diciassettesimo secolo. Gli esperimenti a cui qui si sta facendo riferimento sono quelli di George Artsrouni e Petr Trojanskij, i quali, nel 1933, registrarono, rispettivamente in Francia e Russia, due brevetti. Il primo consisteva in una macchina in grado di svolgere le operazioni di un dizionario meccanico multilingua mentre tra le proposte avanzate da Trojanskij già figurava la possibilità di decodificare e disambiguare regole grammaticali attraverso un processo comprendente una prima fase di analisi logica del testo di partenza, una fase di individuazione e riconoscimento delle forme e strutture equivalenti nel testo di arrivo e un'ultima fase di conversione dell'output nella lingua di arrivo (Hutchins, 2014).

Di lì a poco, quello della traduzione automatica divenne un concetto universalmente compreso dai più e nel luglio del 1949 William Weaver accennò, in una lettera indirizzata a Norbert Wiener, il padre della cibernetica, la reale possibilità che un computer potesse tradurre da una lingua naturale all'altra. Il memorandum di Weaver è oggi considerato lo scritto fondante la teoria sulla traduzione automatica, alla luce delle sfide e degli obiettivi in esso proposti, mentre Weaver è considerato uno dei padri fondatori della traduzione automatica (*Id*, 2001: 9).

Ciononostante, quando le prime ricerche sulla machine translation (MT) iniziarono a prendere piede, i mezzi erano scarsi e il supporto offerto da altre branche della linguistica vicine alla traduzione automatica era pressoché nullo. Di conseguenza, se all'inizio degli anni Cinquanta le aspettative erano alte e si riteneva che importanti scoperte e progressi nel campo fossero ormai a portata di mano, non appena la complessità dei problemi di natura linguistica divenne sempre più evidente

e i ricercatori si scontrarono con quella che fu definita una barriera semantica, agli inizi degli anni Sessanta subentrarono sconforto e disillusione verso quello che cominciava a essere inevitabilmente percepito come un nobile ma irrealizzabile progetto.

Sarebbe ingenuo quindi pensare che, nel corso della sua evoluzione, la MT non abbia mai subito battute di arresto, e ancor di più se si pensa allo scetticismo con cui ancora oggi molti vi si rivolgono. Notabile è quanto, al riguardo, accadde nel 1964. Sebbene le ricerche proseguissero e i primi sistemi funzionanti in grado di produrre traduzioni di bassa qualità venissero installati, nel 1964 il governo degli Stati Uniti incaricò la National Science Foundation di istituire l'Automatic Language Processing Advisory Committee (ALPAC) per esaminare la situazione in cui la traduzione automatica versava al momento. Nel resoconto del 1966, ALPAC sentenziò che la traduzione automatica era lenta, meno accurata e più dispendiosa della traduzione umana e che una sua immediata o quantomeno predicibile utilità era al momento assente (*Id*, 2006: 377). Il resoconto suggeriva, inoltre, che non si investisse più nella ricerca sulla machine translation ma che anzi si sviluppessero strumenti, sì automatici, ma di solo ausilio ai traduttori, come per esempio dizionari automatici. Il motivo per cui ALPAC rifiutò l'idea di un tipo di traduzione automatica era il fatto che questa richiedesse sempre e comunque una fase di post-editing, senza tuttavia considerare che quello della revisione è un passaggio imprescindibile anche nella traduzione condotta da individui. Sebbene fosse considerato poco approfondito e fondamentalmente non obiettivo, il report di ALPAC ebbe degli effetti profondi sulla ricerca sulla MT, segnandone di fatto la fine negli Stati Uniti per almeno più di un decennio, così come altrove, dove mezzi e fondi destinati alla ricerca erano ancora minori.

Perché la ricerca riprenda a fare progressi si dovranno attendere quindi gli anni Settanta e Ottanta. Di grande importanza saranno ad esempio le prime installazioni di Systran, di cui la versione più vecchia fu quella Russo-Inglese installata allo USAF Foreign Technology Division in Ohio nel 1970, seguita nel 1976 dalla versione Inglese-Francese installata dall'allora Commissione delle Comunità europee. Nel corso degli anni, il suo sistema fu poi installato in numerose organizzazioni governative come la NATO e importanti compagnie di rilievo quali la General

Motors. I primi anni Ottanta vedranno invece la nascita del sistema sviluppato dalla Logos Corporation, inizialmente progettato per tradurre manuali di aircraft negli anni settanta, mentre sul finire degli anni Ottanta comparirà sul mercato il sistema tedesco-inglese METAL. Sistemi come questi vennero progettati per essere utilizzati con testi di carattere generale, nonostante i loro dizionari fossero adattati a domini specifici differenti. Cionondimeno, negli anni, fiorì anche la progettazione di sistemi destinati a testi di carattere specifico e funzionali a un dominio particolare.

I progressi compiuti dalla ricerca faranno sì che nel corso degli anni sempre più traduttori diventino familiari con l'idea che i computer possano essere di aiuto al loro lavoro, sebbene non ancora soddisfatti della qualità dell'output che la MT era in grado di restituire. Già allora si stava delineando quella che sarebbe e che tuttora è una corrente di pensiero che vede nella traduzione automatica un valido aiuto, purché il traduttore sia padrone dei processi e non dipenda in tutto e per tutto dai sistemi automatici (*Id*, 2010: 10). In risposta a tali esigenze, negli anni Novanta furono molteplici i CAT tool sviluppati per poter fornire diversi strumenti, tra cui ovviamente i più noti sono termbases, concordance search e translation memory, nonché machine translation. Oggi sono molteplici i fornitori di CAT tool: per nominarne solo alcuni, SDL Trados di SDL, Transit di STAR AG, Déjà Vu di Atril e Wordfast di Champollion.

Sebbene in questa sede verrà posta maggiore attenzione sulla statistical machine translation (SMT), vale la pena ricordare che se è vero che, come si avrà modo di vedere, dagli anni Novanta in poi le ricerche si concentreranno su approcci basati su corpora, fino alla fine degli anni Ottanta la *rule-based machine translation* dominò nel panorama della ricerca sulla MT, con il conseguente proliferare quindi di molteplici regole linguistiche di tipo sintattico, lessicale, morfologico e così via.

Concluso questo brevissimo excursus sulla genesi della machine translation, sembra ora quantomeno opportuno concentrarsi su quello che oggi è l'approccio più diffuso alla machine translation, ovvero la SMT, tuttora in grado di catalizzare su di sé maggior parte della ricerca odierna. Questa analisi da vicino della statistical machine translation permetterà di comprenderne a pieno il suo funzionamento, così da poter fornire le basi per uno studio accurato degli ultimissimi programmi di MT e di quelle che sono le ultimissime tendenze nel mercato della traduzione.

1.2 LA STATISTICAL MACHINE TRANSLATION: COS'È E COME SI CREA UN SISTEMA DI TRADUZIONE AUTOMATICA SMT

Come precedentemente detto, tra gli anni Settanta e Ottanta la ricerca si concentrò sulla SMT quale valida alternativa all'allora dominante *rule-based translation*. La base su cui si fonda la SMT è l'allineamento di frasi, gruppi di parole e singole parole in testi paralleli e il calcolo della probabilità con cui a una data parola in una frase di una lingua corrispondano una o più parole nella frase tradotta e allineata in un'altra lingua. L'approccio che la caratterizza è quindi quello che prevede l'utilizzo di metodi d'apprendimento automatico¹ che consentano a un computer di restituire una traduzione fedele e di qualità. Ciò si ottiene applicando quello che viene definito algoritmo d'apprendimento a un numero cospicuo di testi precedentemente tradotti, meglio noti come corpora paralleli. Così facendo, la macchina è in grado di tradurre frasi mai tradotte prima. Appare evidente, quindi, quanto quantità, qualità e dominio dei testi siano di cruciale importanza per questi sistemi, la cui accuratezza dipende enormemente da questi fattori, sebbene, come si avrà di mostrare successivamente, vi siano testi per cui una traduzione di qualità inferiore si dimostra comunque utile e fruibile.

Prima di scendere nei dettagli, vale tuttavia la pena di riflettere brevemente sui fattori che hanno portato a questa non indifferente crescita dell'interesse verso la SMT. Per quanto concerne la diffusione e il reperimento di informazioni nella propria lingua o in una diversa dalla propria, è opportuno fare riferimento alla distinzione fornitaci da Lopez (2008: 2-3).² Egli distingue tra disseminazione di informazioni e assimilazione di informazioni: la prima concerne tutti quei consumatori interessati nella disseminazione di informazioni in più lingue; la seconda riguarda tutti quegli utenti interessati nell'assimilazione di informazioni in una lingua diversa dalla propria. Appare chiaro che la crescita di Internet ha reso

¹ L'apprendimento automatico, noto in letteratura come *machine learning*, consiste nel migliorare le prestazioni di un programma di computer consentendo a questo di imparare a riconoscere automaticamente e apprendere attraverso l'esperienza e l'elaborazione di dati. Per una descrizione più approfondita e dettagliata del *machine learning* si veda: Mitchell, T. (1997), *Machine Learning*, Boston: McGraw Hill.

² Data la completezza ed esaustività del saggio di Lopez, nel corso di questo capitolo si farà spesso riferimento alla teorie da lui esposte, motivo per cui, per una maggiore comprensione dei temi qui trattati, si consiglia la lettura integrale del suo lavoro, il cui riferimento bibliografico integrale verrà fornito in bibliografia.

possibile entrambi i fenomeni. Organizzazioni governative come l'Unione Europea, la quale ha fatto del multilinguismo la base del suo programma politico, sono ovviamente interessate alla disseminazione di informazioni in più lingue e ciò consente ai ricercatori della SMT di accedere a un ingente numero di testi governativi e notizie bilingui, il cui numero è in costante crescita, essendo questi il frutto di un'attività giornaliera. Di conseguenza, negli ultimi tempi la quantità di fondi stanziati da organi governativi per la ricerca sulla traduzione automatica è aumentata considerevolmente. D'altro canto, Internet ha reso possibile per i fruitori di informazioni occasionali di accedere con relativa facilità a una vasta gamma di informazioni in lingua straniera e tale interesse si riflette in una maggiore richiesta e diffusione di sistemi di traduzione automatica. Ovviamente, i progressi fatti nel campo dell'informatica hanno facilitato la ricerca sulla SMT e, non da ultimo, la spinta a progettare e sviluppare sistemi di traduzione automatica e parametri di valutazione della traduzione automatica ha nutrito la competizione tra gruppi di ricerca, il cui risultato è un circolo virtuoso che ha permesso alla ricerca di progredire. Infine, come si vedrà più avanti, negli ultimi anni sono diventati molti i progetti il cui scopo è la progettazione di programmi di SMT open-source, con il fine di abbattere le barriere e facilitare l'accesso di un vasto numero di utenti alla traduzione automatica.

Come si è avuto modo di accennare poc'anzi, il meccanismo alla base della SMT è un tipo di equivalenza traduttiva che può essere ridotto al livello della singola parola. Chiaramente, la traduzione parola per parola pone un problema piuttosto ostico da superare, ovvero quello dell'ambiguità. Non è infrequente, infatti, che differenti traduzioni di una stessa parola possano avere significati diversi che solo il contesto può aiutare a disambiguare. È quindi opportuno che un programma di traduzione automatica sia in grado, di volta in volta, di scegliere, tra le possibili opzioni, quella più corretta in base al contesto. Inoltre, i traduttori equivalenti raramente occorreranno nello stesso ordine sia nel source text che nel target text. Pertanto, sarà necessario disporre di un meccanismo in grado di riordinare l'output restituitoci tramite machine translation, meccanismo che in genere dipende dalla struttura sintattica della lingua d'arrivo. Fondamentalmente, quindi, la traduzione automatica si può ridurre a una sequenza di passaggi che permettono di ricavare un

testo target da un testo source, il quale non è altro che il prodotto di traduzione parola per parola e riordino dei traducanti ottenuti. Ciò detto, è chiaro che a ogni passaggio corrispondono una serie di ostacoli che un sistema MT funzionante dovrà essere in grado di superare, seguendo una serie di regole e meccanismi impartitigli al momento della sua creazione. Generalmente, quindi, un sistema MT allo stato dell'arte deve essere in grado di far fronte a quattro tipologie di problemi, elencati qui di seguito:

- il primo problema da affrontare è impostare quella serie di passaggi che permettono di individuare la corretta traduzione di un dato testo di partenza all'interno di un numero piuttosto elevato di testi di arrivo. Per fare ciò, ci si serve di un modello di equivalenza traduttiva;
- successivamente, sarà necessario dotare il modello di strumenti, ovvero parametri, che gli consentano di assegnare un punteggio a ogni coppia di segmenti source e target, al fine di disambiguare ogni possibile ostacolo;
- il passo successivo consiste nell'attribuire dei valori ai suddetti parametri in quel procedimento che prende il nome di stima dei parametri;
- infine, sarà necessario cercare e scegliere la traduzione il cui punteggio è il più alto secondo il modello. Tale passaggio prende il nome di decodifica.

Sebbene questa non sia la sede per studiare approfonditamente tutti i meccanismi che portano alla creazione di programma di traduzione automatica, vale ora la pena di esplorare in maniera chiara e sintetica i componenti fondamentali e i passaggi appena illustrati.

1.2.1 CREARE UN SISTEMA DI TRADUZIONE SMT I: MODELLI DI EQUIVALENZA TRADUTTIVA

Come espresso in precedenza, alla base della SMT vi è l'idea che un documento possa essere tradotto calcolando la probabilità che un dato segmento target sia l'esatta traduzione di un dato segmento source. Per consentire un'operazione di questo tipo è necessario disporre pertanto di un modello traduttivo che sia indice della probabilità che uno specifico segmento target sia la traduzione di un altrettanto

specifico segmento source e un modello linguistico che sia indice della probabilità di vedere restituito quel dato segmento target (Brown *et al.*, 1990).

Un modello non è altro che l'insieme di tutte le regole che, di volta in volta, un sistema MT applica per restituire un testo target a partire da un testo source. Fondamentalmente, i modelli utilizzati sono due.

Il primo è quello che viene definito traduttore a stati finiti, ovvero un automa a stati finiti³ che permette di modellare sistemi come quelli presi qui in esame. Esistono due modelli di questo tipo: *word-based* o *phrase-based*. Il modello *word-based*, così come il nome suggerisce, consta della traduzione delle singole parole e del successivo passaggio di riposizionamento delle stesse. A garantire il corretto funzionamento di questo modello vi è un valore che prende il nome di valore di fertilità, il quale indica il numero di parole target generabili a partire da una parola source. Tuttavia, come è ovvio, poiché una parola source può essere allineata a un numero indefinito di parole target, la fase di riposizionamento mostra le difficoltà intrinseche del modello *word-based*, il quale prevede invece un allineamento più lineare. Infatti, i modelli di questo tipo restituiscono traduzioni meno accurate, soprattutto quando si traduce tra lingue il cui ordine sintattico differisce in maniera consistente. Il modello *phrase-based*, d'altro canto, affronta questo problema. In questo modello, in cui il valore di fertilità è assente, l'unità minima non è la parola ma bensì una sequenza di parole che può essere assimilabile alla nostra concezione di sintagma. Ogni sintagma source è qui associato a un sintagma target. Infine, i sintagmi tradotti vengono riordinati nel prodotto finale. A differenza del modello *word-based*, questo modello restituisce traduzioni migliori e anche la fase di riposizionamento crea meno problemi, a patto che, riordinando le varie frasi, la distanza da coprire non sia troppo elevata.

Il secondo tipo di modello prende il nome di grammatica non contestuale o *context-free*. Le grammatiche *context-free* costituiscono le basi per la sintassi di molti linguaggi di programmazione e all'interno della gerarchia di Chomsky delle

³ Un automa a stati finiti è un modello che “descrive un comportamento in cui esiste un numero finito di stati e varie azioni che provocano la transizione da uno stato a un altro” (Beri, 2007: 165).

grammatiche formali occupano il secondo livello.⁴ Il grande vantaggio apportato da questo modello è la stretta relazione che lo lega a una data tipologia di rappresentazione sintattica. Inoltre, tali modelli, nel processo di riposizionamento della frase, sono in grado di coprire distanze maggiormente elevate del modello a stati finiti con uno sforzo relativamente minore. Le tipologie che vengono qui brevemente prese in considerazione sono le *bracketing grammars* e la *traduzione guidata dalla sintassi*. Le *bracketing grammars* fanno uso di alberi sintattici o parentesi quadre per ordinare le stringhe in tutte le loro possibili rappresentazioni, ovviando, come accennato poc'anzi, al problema della distanza. Tuttavia, esistono determinati meccanismi di riposizionamento che neanche le *bracketing grammars* sono in grado di eseguire. L'applicazione della traduzione guidata dalla sintassi invece consente di incorporare informazioni di tipo linguistico e sintattico alle grammatiche non contestuali al fine di rispettare le regole sintattiche della lingua di arrivo nella fase di riordino. Cionondimeno, questa resta solo un'ipotesi, parzialmente smentita da prove empiriche che hanno dimostrato come la sintassi sia sì rispettata entro certi limiti, evidenziando però eccezioni che ricorrono in maniera sistematica.

Alla luce dell'evidente punto debole di questi modelli, senza entrare nei dettagli, è tuttavia opportuno fare riferimento all'esistenza di un terzo tipo di modello, il quale combina la struttura sintattica con i modelli *phrase-based*, i quali come si è visto sono in grado di risolvere i problemi caratterizzanti i modelli *word-based*.

1.2.2 CREARE UN SISTEMA DI TRADUZIONE SMT II: PARAMETRIZZAZIONE

Impostare un modello di equivalenza traduttiva, tuttavia, non è sufficiente a permetterci di scegliere, tra le innumerevoli possibili frasi target, quella corretta, poiché l'ambiguità della lingua rende necessario l'impiego di un meccanismo che

⁴ Dato il considerevole impatto che la gerarchia di Chomsky ha avuto sulla concezione dei linguaggi formali e delle scienze matematiche, informatiche e linguistiche, nell'ottica di una maggiore comprensione dei modelli qui analizzati, si vedano i due scritti fondanti di tale teoria: Chomsky, N., (1956) "Three models for the description of language", *IRE Transactions on Information Theory*, 2, 113-124 e Chomsky, N., (1959) "On certain formal properties of grammars", *Information and Control*, 1, 91-112

consenta al sistema di scegliere la coppia source-target adatta. Tale meccanismo prende il nome di parametrizzazione.

Come già ribadito, le difficoltà intrinseche che scaturiscono dall'inevitabile ambiguità della lingua sono dettate dal fatto che il compito di un sistema MT è quello di processare diverse tipologie di fenomeni linguistici specifici conservando al tempo stesso tutta una serie di informazioni su lingua source e lingua target che sono indipendenti dalla lingua. Lo scopo è quindi quello di dotare la macchina di una serie di principi universali ritenuti tali in tutte le lingue e di una serie di parametri che di volta in volta consentano alla macchina di determinare se tali principi vengono realizzati o meno nell'input e output di ogni lingua (Dorr, 1992). A tale proposito, quindi, l'impiego di parametri apporta vantaggi non indifferenti, quali ad esempio la possibilità di rappresentare tutta una serie di competenze linguistiche specifiche indipendenti da quelle dei principi universali e semplificare notevolmente la rappresentazione di regole grammatiche, consentendo ciononostante la descrizione di una vasta gamma di fenomeni linguistici complessi (*Id*, 1993). A tale scopo, per chiarire il funzionamento di questo procedimento ci si servirà dell'esempio di UNITRAN, un sistema di traduzione automatica che traduce inglese, spagnolo e tedesco, oggetto di studio e analisi della succitata Dorr (1987). La caratteristica principale di UNITRAN, acronimo di UNiversal TRANslator, è la sua capacità di operare in maniera uniforme attraverso le lingue, considerando al tempo stesso le specificità di ogni lingua. È chiaro che un processo di questo tipo è tutt'altro che semplice, alla luce dell'enorme vastità di fenomeni linguistici che possono ricorrere in una coppia di frasi. Si esamini, per l'appunto, il seguente esempio da Dorr (1990: 13):

- (1) Juan forzó la entrada al cuarto
John broke into the room

La frase source diverge da quella target sia dal punto di vista lessicale che da quello sintattico. Infatti, dal punto di vista sintattico, le due frasi realizzano l'argomento del verbo diversamente, poiché alla parola "room" corrispondono sia "entrada" che "cuarto". Dal punto di vista lessicale, invece, in inglese l'azione del

verbo viene descritta con il verbo “broke” mentre in spagnolo dal sintagma verbale “forzar la entrada”. Per risolvere tali divergenze il sistema opera su due livelli, dei quali uno è quello lessicale-semantic, astratto e distinto dall’altro livello, quello sintattico. La traduzione non è altro che il risultato dell’interazione tra questi due livelli. Tale interazione è ovviamente resa difficile dalle diverse informazioni di tipo linguistico e non che il sistema deve elaborare contemporaneamente. Ad esempio, come si evince dal seguente caso, tratto sempre dal lavoro di Dorr (*ibid.*) sulla base delle informazioni sulla lingua inglese elaborate dal sistema, il risultato della traduzione di una frase spagnola in inglese è piuttosto predicibile:

(2) Qué golpeo Juan

In base alla conoscenza che il sistema ha della lingua inglese, agente e paziente verrebbero quindi invertiti e tradotti come segue:

(3) What hit John

Per evitare che ciò accade è quindi opportuno, come già detto in precedenza, che il sistema di traduzione disponga di conoscenze che sono solo specifiche della lingua e di conoscenze indipendenti da essa, derivabili quindi dal contesto, per esempio, atte a interpretare correttamente la frase. A tale scopo, pertanto, prendendo sempre a modello UNITRAN, è possibile osservare un approccio traduttivo fondato su principi che hanno valore universale in tutte le lingue e parametri ad essi associati che consentono di elaborare i tratti distintivi di suddette lingue. Ad esempio, l’assenza o presenza del soggetto in una frase costituisce uno dei principi a cui il sistema deve fare riferimento di volta in volta, attraverso l’impiego di un parametro associato al principio del soggetto sottinteso, il quale sarà applicabile per lingue quali lo spagnolo e l’italiano e non applicabile a lingue quali l’inglese. È infatti questo il caso dei succitati esempi (2) e (3): il parametro consente al sistema di riconoscere lo spagnolo quale lingua in grado di reggere un soggetto sottinteso e, di conseguenza, avere un ordine dei componenti della frase meno vincolante, tale da ammettere la presenza di un soggetto in posizione post-verbale, il quale sarà quindi interpretato come soggetto

della frase, anziché oggetto, anche nell'equivalente inglese. Nel sistema UNITRAN tale analisi è possibile grazie a un semplice menù tramite il quale impostare i parametri del modello traduttivo, come mostrato in figura 1.1.

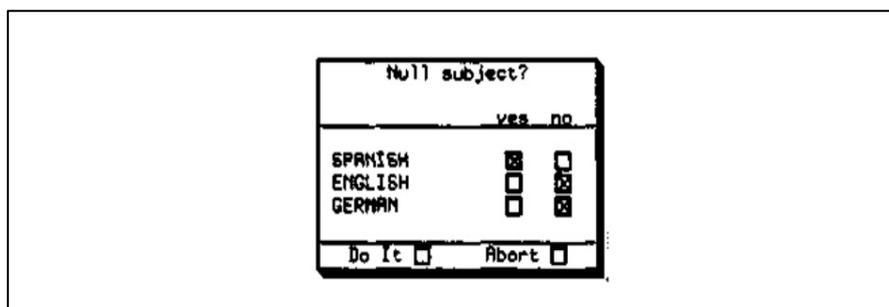


Figura 1.1 Schermata parametro *Soggetto sottinteso* nel sistema UNITRAN. (Dorr, 1990: 15)

1.2.3 CREARE UN SISTEMA DI TRADUZIONE SMT III: STIMA DEI PARAMETRI

Una volta definito il modello di equivalenza traduttiva e ad esso applicati i parametri, il passo successivo sarà stimare i suddetti parametri, assegnando ad essi dei valori, in modo tale da prevedere l'esatta probabilità con cui essi occorreranno nel testo. Questo passaggio risulta fondamentale per fornire al sistema informazioni e strumenti necessari ad applicare di volta in volta il giusto parametro. Tale procedimento è realizzabile tramite l'impiego di un corpus parallelo grazie al quale la macchina può imparare i valori dei parametri, stimando quindi la probabilità che quel parametro sia quello corretto e che quindi la traduzione sia quanto più vicina possibile a quella esatta (Lopez, *op. cit.*).

Dal momento che, come si è già avuto modo di vedere, i modelli di equivalenza traduttiva possono adottare tanto un approccio traduttivo parola per parola, quanto un approccio frase per frase, è opportuno distinguere e precisare metodologie e problematiche scaturenti dalla valutazione dei parametri dei due differenti modelli. In entrambi i casi è possibile ricorrere ad allineamenti del corpus utilizzato per ogni modello. Nel caso del modello parola per parola, tuttavia, tale compito è reso difficile dal fatto che gli allineamenti forniscono coppie di frasi in cui, sebbene sia vero che parole equivalenti cooccorrono, non è dato conoscere l'esatto numero di volte in cui tali parole sono effettivamente allineate le une alle altre. Al fine di

adottare il metodo della massima verosimiglianza (ivi: 23). Senza entrare troppo nel dettaglio e rischiare di confondere il lettore con meccanismi e principi che sconfinano nelle scienze statistiche e nell'informatica, si preciserà solo che tale procedimento prende il nome di training dei parametri e dal momento che tale training non è altro che una tipologia di apprendimento automatico, disponendo degli input, dei corrispondenti output e dei parametri, il modello della massima verosimiglianza menzionato poc'anzi permette di stimare i parametri del modello di traduzione automatica.⁵

1.2.4 CREARE UN SISTEMA DI TRADUZIONE SMT IV: DECODING

Ora che si dispone di un modello, di cui sono stati stabiliti e stimati i parametri, è ora finalmente giunto il momento di tradurre le frasi source, ovvero operare una decodifica. Questo passaggio conclusivo può avvenire in diversi modi. Il più generico è quello descritto da Wang e Waibel (1997) e Callison-Burch (2007) e prende il nome di stack decoding. Tale metodo prevede la ricerca e organizzazione delle migliori ipotesi di traduzione in stacks, ovvero pile, a partire da sequenze di parole costruite partendo da sinistra verso destra, seguendo l'ordine della lingua target, come mostrato in figura 1.3:

⁵ Per quanto non sia queste la sede per approfondire concetti e temi che trovano la loro origine nelle scienze statistiche, per evitare che l'applicazione del metodo della massima verosimiglianza applicato alla stima dei parametri rimanga un concetto alieno al lettore, si veda: Berger, A.L., Della Pietra, V.J. e S.A. Della Pietra (1996) "A maximum entropy approach to natural language processing". *Computational linguistics*, 22(1), 39-71.

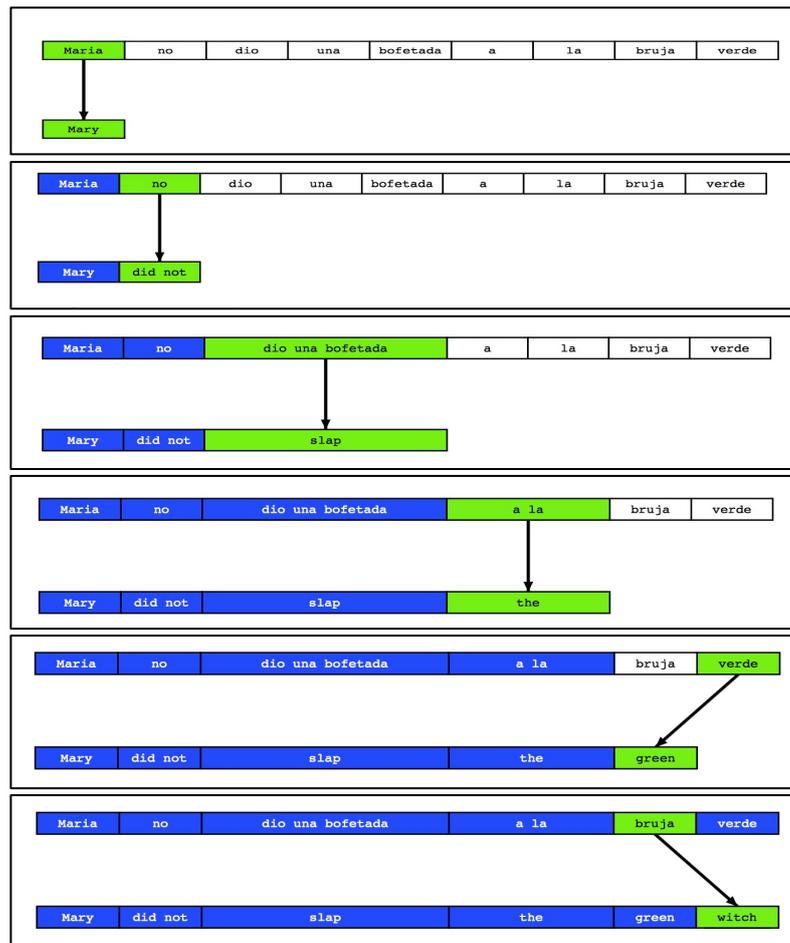


Figura 1.3 Processo di costruzione della frase tradotta. La frase viene suddivisa in unità grammaticali e di ognuna viene ricercato il traduttore esatto; una volta tradotta, l'unità grammaticale viene contrassegnata come tradotta (Callison-Burch, 2007: 5-11).

Tale processo di espansione delle ipotesi prevede che si parta da un'ipotesi vuota dalla quale iniziare la pila; in seguito, scelta l'ipotesi traduttiva migliore, la prima parola source viene tradotta e l'ipotesi, estratta dalla pila, viene isolata e aggiunta in un'altra e così via. Le ipotesi espanse vengono organizzate e aggiunte in pile in base a stessa parola source tradotta, stesso numero di parole source tradotte e stesso numero di parole target restituite (Wang and Waibel, *op. cit.*). In base a questo meccanismo, le ipotesi meno probabili vengono scartate e solo quelle con la probabilità più alta vengono estratte dalla pila, come mostra Callison-Burch (*op. cit.*, 26) nella figura a pagina seguente:

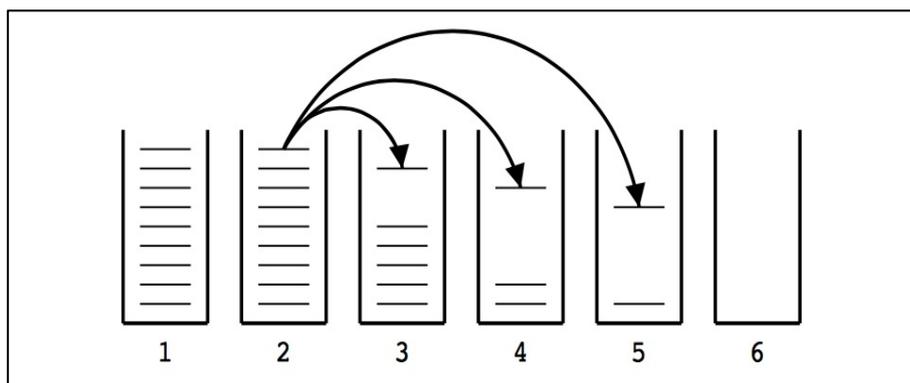


Figura 1.4 Rappresentazione semplicistica ma efficace del processo di organizzazione delle ipotesi espanse in pile (*hypotesis stacks*).

Comparare una o più ipotesi, al fine di scegliere quella corretta, è un procedimento relativamente semplice poiché, confrontando le ipotesi traducenti lo stesso numero di parole, il sistema sceglierà l'ipotesi che traduce quelle porzioni della frase più semplici e scarterà le altre.

Alternativamente, un metodo di decoding adottato è quello che prende il nome di *greedy decoding* (Germann, 2001). Tale metodo di decoding consiste nel costruire la frase tradotta allineando a ogni parola source la parola target più probabile. Ottenuto questo allineamento, il motore di decoding cerca di migliorarlo trovando allineamenti le cui probabilità sono maggiori, tramite una serie di operazioni, che possono consistere nella traduzione di una o due parole, la traduzione di una parola e l'inserimento di una nuova, la rimozione di parole non allineabili e così via, come mostrato in figura 1.5. a pagina seguente.

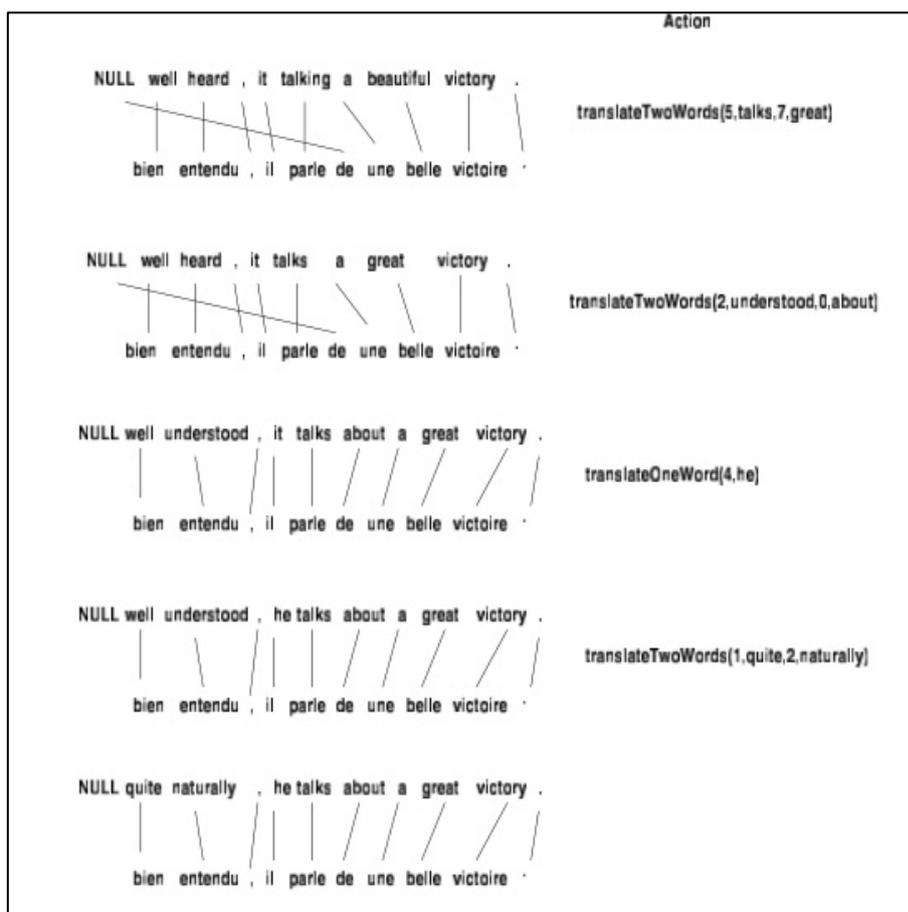


Figura 1.5 Schematizzazione del processo di *greedy decoding*. A partire dalla frase “Bien entendu, il parle de une belle victoire” il processo di *greedy decoding* produce, attraverso una serie di passaggi volti a migliorare l’allineamento iniziale, la frase “Quite naturally, he talks about a great victory.” (Germann, 2001: 5).

In conclusione, quindi, questi quattro passaggi appena elencati e descritti permettono di generare un sistema di traduzione automatica. Tuttavia, come appare chiaro, tali sistemi non sono perfetti e, intrinsecamente, di volta in volta possono incappare in difficoltà e restituire pertanto traduzioni di bassa qualità. È per questo motivo che, di pari passo alla ricerca e progettazione di sistemi di traduzione automatica, diversi studi sono stati concentrati sull’impostazione di parametri e metodologie di valutazione dell’output della machine translation. A seguire, si elencheranno e descriveranno le proprietà e caratteristiche dei più usati e conosciuti.

1.3 VALUTARE L'OUTPUT DI UN SISTEMA DI TRADUZIONE AUTOMATICA

Tradizionalmente, si considerano accettabili quali metodologie di valutazione dell'output della MT tutte quelle condotte tramite l'esaminazione da parte di traduttori professionisti, incaricati di valutare due parametri fondamentali per giudicare correttamente la qualità di una traduzione: l'adeguatezza tramite la quale la traduzione comunica lo stesso messaggio del testo di partenza e la scorrevolezza del testo di arrivo. Tuttavia, se da un lato queste si rivelino essere le metodologie di valutazioni migliori e più affidabili, dall'altro lato, sfortunatamente, sono anche quelle che richiedono più tempo e denaro. Dal momento che valutazioni di questo tipo possono richiedere settimane o persino mesi, a fronte dell'esigenza di monitorare quotidianamente un sistema di traduzione automatica, la valutazione condotta da traduttori professionisti si rivela oltremodo inadeguata. Di conseguenza, diviene necessario sviluppare metodi automatici di valutazione in grado di restituire gli stessi risultati che si avrebbero se tali valutazioni fossero condotte da individui in carne e ossa (Lopez, *op. cit.*).

L'elemento principale condiviso da ogni metodo automatico di valutazione è l'utilizzo di un numero variabile di frasi delle quali sia disponibile una traduzione prodotta da traduttori professionisti. Tali frasi tradotte vengono definite *reference translation*. L'idea alla base di tali valutazioni è che un sistema MT è tanto buono quanto più l'output da esso prodotto si avvicina a quello di un traduttore professionista. Affinché una valutazione di questo tipo possa considerarsi efficace e utile allo scopo, i criteri da soddisfare sono diversi. Ovviamente, come già accennato, un requisito fondamentale è che la valutazione automatica sia strettamente correlabile con quella umana. Inoltre, tali valutazioni devono essere sensibili anche a minime differenze nella qualità dell'output tra sistemi differenti. Infine, i risultati di tali sistemi dovrebbero essere concordi tra loro quando si applica uno stesso sistema di MT su testi simili ed assicurare un certo grado di affidabilità quando sistemi MT diversi restituiscono gli stessi risultati.

Di seguito, si analizzeranno brevemente le caratteristiche di alcuni sistemi automatici di valutazione dell'output della machine translation, evidenziando tanto gli aspetti condivisi tra gli stessi quanto le prese di distanza che differenziano gli uni dagli altri.

1.3.1 BLEU: BILINGUAL EVALUATION UNDERSTUDY

Come precedentemente asserito, la regola aurea per la valutazione di un sistema MT è la relativa vicinanza, se non sovrapposizione, tra il prodotto della machine translation e quello di un traduttore professionista. Questa è l'idea alla base del metodo automatico di valutazione BLEU, acronimo di *Bilingual Evaluation Understudy*. Modellato su un altro metodo denominato WER (*word error rate*), il quale somma il numero di inserimenti, cancellazioni e sostituzioni necessari affinché la machine translation sia equivalente alla reference, BLEU prende da quest'ultimo i suoi vantaggi e ne affronta invece i punti deboli: BLEU non si limita a considerare i match tra singole parole ma tra *n-gram*, dove *n* può equivalere a un numero massimo variabile di combinazione di parole. Inoltre, laddove WER assegna una penalizzazione a una data parola corretta ma in una posizione sbagliata, BLEU, in caso di posizionamento errato di una parola tradotta correttamente, assegna un punteggio differente dalla cancellazione o sostituzione della stessa (Och, 2003).

In seguito, una volta assegnati i punteggi a ogni singola frase tradotta, si effettua una stima della qualità globale della traduzione operando una media di tutti i punteggi ottenuti. Al raggiungimento del punteggio per ogni singola frase concorrono diversi fattori, tra cui eventuali penalizzazioni per quanto concerne la lunghezza della frase. Infatti, idealmente, una frase *machine translated* non dovrebbe essere né troppo lunga né troppo corta. Inoltre è prevista una penalizzazione per ogni parola che occorre più frequentemente delle volte che occorre nella reference translation.

I punteggi assegnati da BLEU sono compresi tra 0 e 1, dove le traduzioni che totalizzeranno 1 saranno veramente poche, ad eccezione di quelle che risultano identiche alle reference translation. Ciò prevede, di fatto, che neanche un traduttore professionista raggiungerà un punteggio equivalente a 1, alla luce del fatto che tale punteggio è anche influenzato dal numero di reference translation utilizzate per ogni frase tradotta tramite MT: maggiore è il loro numero, maggiore sarà il punteggio (Papineni *et al.*, 2002).

1.3.2 NIST: NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY

NIST, progettato dal *National Institute of Standards and Technology*, è un metodo di valutazione automatico dell'output della machine translation pensato per affrontare e risolvere le problematiche scaturenti da BLEU, con il quale ciononostante condivide alcune caratteristiche. Una differenza fondamentale rispetto alla metodologia di valutazione appena analizzata è l'importanza che NIST assegna al grado di informatività di un dato n-gram: qualora la traduzione di tale n-gram sia corretta, l'algoritmo di NIST calcola quanto tale n-gram sia comune o meno e la sua frequenza all'interno del corpus oggetto di studio; il risultato così ottenuto consente di riportare il punteggio dato a ogni singolo n-gram in base al suo peso in termini di informatività (Doddington, 2002). Inoltre, a differenza di BLEU, variazioni minime nella lunghezza della frase hanno un impatto minore sul punteggio della stessa.

1.3.3 METEOR: METRIC FOR EVALUATION OF TRANSLATION WITH EXPLICIT ORDERING

Progettato a sua volta per bypassare alcuni dei problemi nati dall'utilizzo di BLEU e non del tutto risolti da NIST, METEOR, acronimo di *Metric for Evaluation of Translation with Explicit Ordering*, introduce caratteristiche inedite alle metodologie di valutazione automatica antecedenti e dimostra di possedere una buona correlazione con le valutazioni condotte da individui, correlazione utile a individuare anche minime differenze tra sistemi MT diversi e tra diverse versioni dello stesso sistema MT (Banerjee, 2005). Inoltre, anche il processo di matching tra l'output della MT e una o più reference translation subisce alcune modifiche al fine consentire una maggiore elasticità ed espanderne le modalità. Oltre al matching di n-gram identici, infatti, anche varianti morfologiche simili, come ad esempio due parole che condividono la stessa radice, costituiscono un match; analogamente, anche la sinonimia costituisce un match e quindi, quando due parole sinonimi l'una dell'altra condividono un set di sinonimi stabilito tramite database esterni, queste danno luogo a un match.

Infine, qualora sia necessario il riposizionamento di parole tra la frase machine translated e la reference translation, al punteggio di ogni singola frase viene assegnata una penalizzazione.

1.3.4 TER: TRANSLATION EDIT RATE

TER, acronimo di *Translation Edit Rate*, è un metodo di valutazione automatica definibile come il minor numero di modifiche necessario affinché l'output della machine translation sia identico a una delle reference translation, diviso il numero medio di parole nella reference translation (Snover *et al.*, 2006). Le modifiche permesse includono inserimenti, cancellazioni, sostituzione di parole o loro sequenze. È importante notare, tuttavia, che tutte le modifiche hanno lo stesso peso e quindi riposizionare una parola avrà lo stesso impatto sul punteggio di una cancellazione, poiché entrambe sono considerabili come modifiche. Considerazioni analoghe valgono per la punteggiatura e l'uso errato di maiuscole e minuscole. Per chiarire al meglio il funzionamento di TER, serviamoci dell'esempio presentatoci da Snover (ivi) a supporto del suo elaborato, dove si presenta una coppia data da reference translation e ipotesi generata dalla MT; le modifiche apportate sono indicate in maiuscolo:

REF: SAUDI ARABIA denied THIS WEEK information published in the
AMERICAN new york times

IPO: THIS WEEK THE SAUDIS denied information published in the new york
times.

Sebbene, dal punto di vista semantico, ipotesi e reference si possano considerare equivalenti, TER non considera questa coppia come un match esatto. Le modifiche necessarie perché si abbia un'equivalenza perfetta sono le seguenti: la posizione di THIS WEEK deve essere modificata, SAUDI ARABIA appare come THE SAUDIS nell'ipotesi e quindi necessita di due sostituzioni e infine AMERICAN non appare nell'ipotesi. Il numero di modifiche necessarie ammonta quindi a quattro, numero

che, diviso per il totale delle parole nella reference (13), restituisce un punteggio TER di 31%.

Sebbene questo metodo tenga in considerazione la reference translation più vicina possibile all'ipotesi, TER non può certo considerare il concetto di equivalenza semantica. Per tale motivo, servendosi dell'ausilio di parlanti della lingua target, si può ricorrere a una versione "umanizzata" di TER, la quale prende il nome di HTER, acronimo di *Human-targeted Translation Edit Rate*. Lo scopo di HTER è quello di trovare la reference translation migliore e più vicina possibile all'ipotesi tra tutte le reference translation che hanno lo stesso significato di quella originale. A tal fine, la procedura prevede che gli editor, a partire da una ipotesi generata da un sistema MT e una o due reference translation, generino una reference definita *targeted*, modificando l'ipotesi oppure la reference translation originale. Infine, la *targeted reference translation* così ottenuta viene utilizzata per elaborare i punteggi di HTER.

2. IL POSTEDITING: ANALISI DEGLI STANDARD PRODUTTIVI E QUALITATIVI

2.1 VALUTARE GLI STANDARD QUALITATIVI DELLA TRADUZIONI AUTOMATICA: PARAMETRI DI MACRO E MICROVALUTAZIONE

Nel capitolo precedente c'è stato modo di esplorare le meccaniche alla base del funzionamento di un sistema MT ed apprezzarne caratteristiche principali, punti di forza e punti deboli. In conclusione a suddetto capitolo, il lettore ha avuto modo di usufruire di una breve visione d'insieme delle principali metodologie di valutazione dell'output di un sistema di traduzione automatica, di cui si è sì lodata la velocità di esecuzione e il costo necessario inferiore, ma di cui non si è di certo potuto dire la stessa cosa in quanto a correlazione con le valutazioni condotte da individui professionisti. La domanda che, pertanto, spesso sorge è questa: a fronte di una spesa minore e di una produttività maggiore, fino a che punto, in virtù della qualità dell'output, la machine translation può essere tale da considerarsi un asset per il traduttore? E in che misura il post-editing, o correzione manuale dell'output della machine translation, può considerarsi la soluzione adatta a colmare suddetto *quality gap* (Green, 2013)? Le risposte a queste e altre domande possono essere molte e non tutte semplici da trovare, alla luce dei numerosi fattori che entrano in gioco, e

ciononostante, per quanto possibile, nel corso di questo capitolo si cercherà proprio di trovare delle risposte quantomeno esaustive ad alcuni di questi interrogativi.

Prima di entrare nel vivo di un discorso così complesso pare tuttavia opportuno spendere alcune parole per sfatare alcune delle errate concezioni su machine translation e post-editing, al fine di poter poi affrontare questa spinosa questione con chiarezza e dati certi alla mano. Nel capitolo introduttivo a *Machine Translation* (1994), Arnold *et al.* elencano in maniera esaustiva e dettagliata alcuni tra i più famosi pregiudizi e false credenze sulla MT, per poi confutarli in maniera ineccepibile. Tra questi, ovviamente il più comune non può che essere quello relativo a una qualità dell'output della machine translation così scarsa da renderne inutile l'impiego. Di fatti, come è lecito aspettarsi, la qualità di molti sistemi MT è sì bassa, ma cionondimeno utile a fornire all'utente una traduzione non esente da errori e incomprensioni ma fruibile tanto da un parlante della lingua target interessato a comprendere le nozioni basilari del testo tradotto quanto da un post-editor, il cui compito sarà quello di revisionare il testo tradotto automaticamente ed elevarne la qualità (ivi: 8). Come verrà espresso più avanti, del resto, quello della qualità è un discorso complesso e sfuggibile, talvolta soggettivo per giunta, e stabilire una metodologia di valutazione trasversale a tutti domini e le tipologie testuali, nonché i sistemi MT, può risultare un'impresa quanto mai ardua, alla luce delle innumerevoli variabili che insieme cooccorrono a delineare una traduzione, quali possono essere la tipologia testuale e il contenuto, oppure lo scopo per cui un dato testo deve essere tradotto, nonché il fine e gli utilizzatori ultimi della traduzione. È inoltre opportuno ricordare che se da un lato vi sono caratteristiche linguistiche intrinseche in un testo tradotto che insieme ne definiscono la qualità, dall'altro, nel mercato odierno, quello della qualità è un concetto che risulta essere profondamente e indissolubilmente legato ai concetti di tempo e costi.⁶

Sullo stesso livello, se non ancor più infondata, è l'idea che la traduzione automatica renderà obsoleta la figura dei traduttori. Ancora una volta la contro risposta a questa falsa credenza risiede nel post-editing: alla luce dei tuttora effettivi limiti degli attuali sistemi MT è quantomeno improbabile che tali sistemi

⁶ Per approfondire il concetto, si veda il webinar "Big Picture View on Quality Evaluation", estratto dal TAUS Quality Evaluation Summit (2013), del quale è stato riadattato parte del contenuto citato. Il link al video integrale è disponibile nella sezione Bibliografia.

sostituiranno ed elimineranno, in un immediato futuro, la figura del traduttore professionista, condizionandone e controllandone il lavoro. Ciò che è auspicabile e quantomeno predicibile, al contrario, è una maggiore implementazione della traduzione automatica, quale strumento di perfezionamento della produttività dell'individuo professionista del settore, poiché attraverso l'automatizzazione di determinati processi diverrà più facile garantire coerenza terminologica in determinati settori (*ibid.*).

Quanto detto finora, tuttavia, avrebbe poco fondamento se prima non si cercasse di definire, per quanto possibile, alcune linee guida utili a discriminare, in maniera quantitativa e qualitativa, tra qualità traduttiva medio-alta e qualità bassa-insufficiente. Come già detto poc'anzi, quello della qualità è un concetto relativo, valutabile entro certi limiti imposti da un grado di soggettività dal quale è impossibile esimersi nel momento in cui si cerca di quantificare fenomeni astratti e concettuali (Bruderer, 1978 cit. in Van Slype, 1979: 32). Generalmente, i parametri che si cerca di includere in una valutazione della qualità traduttiva sono la morfologia, la sintassi, il contenuto, lo stile e la terminologia, a seconda che si valuti in presenza o in assenza del testo source. Nel primo caso, qualora si disponga del testo di origine, sarà importante valutare se, sulla base dello scopo della traduzione, il contenuto del testo originale è stato riprodotto in maniera fedele e nel rispetto delle intenzioni dell'autore; sulla base della tipologia testuale, d'altro canto, sarà di fondamentale importanza verificare che tutte le informazioni contenute nel testo di partenza siano state riportate nel testo di arrivo, accertandosi che il testo sia costruito in modo tale da garantire una comunicazione efficace. Nel secondo caso, ovvero quando non si dispone del testo di partenza e si valuta quindi la traduzione come testo indipendente, sarà più opportuno valutare gli aspetti più formali del testo, quali utilizzo di una terminologia appropriata alla tipologia testuale e al contesto comunicativo, l'utilizzo di uno stile e di costruzioni ed espressioni idiomatiche consone, le modalità entro cui si realizza lo scopo illocutivo del testo, nonché, ovviamente, la correttezza grammaticale e sintattica e la coerenza testuale (ivi: 36).

Ciò detto, appare evidente che la valutazione della qualità traduttiva, applicata tanto alla traduzione convenzionale quanto a quella automatica, sia un compito tutt'altro che facile e immediato, al punto che impostare una categorizzazione o

tassonomia dei parametri da prendere in considerazione nella valutazione sembra la scelta migliore o quantomeno vantaggiosa in termini di praticità e applicabilità, sebbene, come si evincerà a breve, ricorrere a categorizzazioni eccessivamente severe e nettamente delineate tra loro possa rivelarsi complicato e dispendioso, se non addirittura controproducente. In tal senso, Van Slype, (ivi: 56) propone una tipologia di valutazione che prende il nome di macrovalutazione, ovvero un'operazione che consiste nel valutare la risposta della traduzione ai differenti requisiti dell'utenza e del contesto lavorativo, giudicando la performance per sé, ovvero scevra da eventuali analisi delle cause che possono avere portato a prestazioni non adeguate. La tassonomia proposta consta di dieci classi, suddivise in quattro gruppi, in base al livello della valutazione affrontato da ciascun gruppo.

<p>Livello cognitivo efficacia della comunicazione e della trasmissione di informazioni</p>	<ul style="list-style-type: none"> • Intelligibilità • Fedeltà • Coerenza • Utilizzabilità • Accettabilità
<p>Livello economico valutazione di tempi e costi</p>	<ul style="list-style-type: none"> • Tempo necessario per la lettura • Tempo necessario per la correzione • Tempo necessario per la traduzione
<p>Livello linguistico conformità al modello linguistico</p>	
<p>Livello operativo valutazione delle operazioni eseguite da un sistema MT</p>	

Tabella 2.1 Livelli e classi del modello della macrovalutazione proposto da Van Slype.
(1979)

Intelligibilità È definibile come la facilità con cui una traduzione può essere compresa da un lettore, nella maniera più chiara possibile. È valutabile, in maniera più o meno soggettiva, giudicando il grado di chiarezza e comprensibilità di un testo tradotto, partendo da un testo perfettamente chiaro e intellegibile, senza alcun tipo di pecca stilistica, a un testo totalmente incomprensibile, neanche dopo un considerevole quantitativo di studio e analisi dedicatogli, passando per testi mediamente intellegibili, la cui comprensione può richiedere una conoscenza approfondita e ben strutturata della materia trattata. È sicuramente uno dei metodi di valutazione più usati, poiché consente di porre direttamente in relazione qualità traduttiva e suo effetto sul lettore. La sua utilità è confermata inoltre dal suo largo utilizzo come criterio di valutazione di *drafts*, ovvero testi sui quali è necessario il lavoro di un post-editor, affinché quest'ultimo ne elevi la qualità generale. Tra i vantaggi apportati dal suo impiego vi è sicuramente l'immediatezza della valutazione poiché il testo viene presentato così come è all'ipotetico valutatore, al quale basta possedere una conoscenza essenziale della materia trattata. D'altro canto, il grado di soggettività a cui è sottoposto questo criterio rappresenta uno svantaggio da tenere in forte considerazione, qualora impiegato.

Fedeltà Consiste nella misurazione del grado di correttezza delle informazioni tradotte dalla lingua source alla lingua target. Una valutazione condotta seguendo questo tipo di criterio deve ovviamente tenere conto tanto di perdite di informazioni dovute all'omissione di parole quanto di interferenze provocate dall'aggiunta di parole e informazioni non presenti nel testo originale. Questo tipo di procedimento, come puntualizzato in ivi: 72, è tutt'altro che facile, dal momento che la veicolazione di informazioni non avviene per mezzo di singole entità contenute in una frase, ma bensì attraverso messaggi o serie di messaggi complessi, dei quali stabilire la relativa importanza non è spesso immediato. Pertanto, sebbene sussista un certo grado di equivalenza nei risultati ottenibili applicando i parametri di intelligibilità e fedeltà, dal momento che un'informazione non intellegibile è per forza di cose assente nel testo originale, il criterio della fedeltà risulta di più difficile utilizzo, dal momento

che un suo impiego richiederebbe una preliminare operazione di decodifica e comprensione dei contenuti di ogni singola frase o, alternativamente, l'impiego di una scala di informatività come quella proposta da Carroll (1966 cit. in ivi: 73), atta ad assegnare, su una scala da 9 a 1, un grado di informatività a ogni frase, un'operazione senz'altro utile ma dispendiosa in termini di costi e tempi.

Coerenza Proposto come criterio di valutazione da Wilks (1978 cit. in ivi: 81), quello della coerenza è un criterio che non trova molti riscontri. Alla sua base vi è l'idea che la qualità di una traduzione possa essere valutata in funzione del suo livello di coerenza. Un tipo di coerenza di questo tipo può essere condotta, pertanto, anche da un valutatore monolingue. Ciò che tuttavia lascia perplessi è il fatto che, come nota lo stesso Wilks, la valutazione della coerenza di un testo tradotto non potrà mai essere assoluta in sé, ma sempre in relazione alla coerenza del testo originale, una condizione che rende quindi comunque necessario l'impiego di valutatori bilingui.

Utilizzabilità È definibile come la possibilità che una traduzione possa essere utilizzata. Le metodologie di valutazione condotte adottando questo criterio sono molteplici e, generalmente, prevedono tutte l'impiego di questionari somministrabili agli utenti ultimi del testo tradotto o, in alternativa, a traduttori professionisti ai quali viene richiesto di predire il fine ultimo della possibile utenza di suddetti testi. Tali questionari possono essere strutturati sulla base di differenti parametri quali assi di variazione linguistica (diatopico, diastratico, diafasico, diamesico, diacronico) dei parlanti della lingua target, tipologia testuale (testi informativi, testi narrativi, testi argomentativi) e tipologia dell'approccio traduttivo (traduzione semantica-formale e traduzione comunicativa-funzionale), come proposto da Juliane House (1977 cit. in ivi: 86-90). Ciononostante, il punto debole di questo metodo risiede nel fatto che le situazioni in cui i valutatori sono posti per rispondere a tali questionari sono del tutto artificiali, poiché solo e solamente l'utenza finale può valutare l'importanza di un testo e la conseguente possibilità per questo di essere utilizzato, senza contare che operazioni di questo tipo si rivelano spesso svantaggiose da un punto di vista economico produttivo.

Accettabilità È equivalente al grado secondo il quale una traduzione può essere considerata accettabile dall'utente finale. Il vantaggio derivante dall'utilizzo di questo criterio è, indubbiamente, il fatto che ad esprimere il giudizio è il committente stesso della traduzione, per il quale, un testo tradotto è, semplicemente, accettabile o non accettabile. Si tratta di un criterio di valutazione relativamente facile da implementare poiché, facendo a meno di stadi valutativi parziali o intermedi, si pone in stretta relazione con lo stadio ultimo del processo traduttivo, ovvero l'accettazione da parte dell'utente del testo, sebbene poi quest'ultimo faccia, per forza di cose, riferimento a criteri precedenti quali, ad esempio, l'intelligibilità o la fedeltà. D'altro canto, data la vasta gamma di variabili che entrano in gioco, quali tipologia testuale, scopo della traduzione e così via, risulta difficile fare di questo un criterio di valutazione universale in grado di restituire risultanti concludenti, se non applicato a vastissimi campioni di utenti e documenti tradotti (ivi: 113).

Tempo necessario per la lettura Tale criterio si dimostra essere efficiente e di facile attuazione, poiché per il suo utilizzo è solo necessario misurare il tempo necessario per leggere un testo tradotto tramite MT, a confronto con il tempo necessario per leggere lo stesso testo tradotto da un traduttore professionista: ogni variazione tra i due dati sarà indicativa del tempo in più necessario alla lettura, quale causa di una qualità traduttiva del sistema MT inferiore. Si tratta inoltre di un criterio di valutazione per nulla dispendioso, poiché conducibile in concomitanza con la valutazione dell'intelligibilità di un testo.

Tempo necessario per la correzione Analogamente a quello precedente, anche suddetto criterio è di facile utilizzo, dal momento che è sufficiente misurare il tempo necessario affinché un revisore o un post-editor attuino le correzioni necessarie, meno invasive nel primo caso e con ogni probabilità maggiori nel secondo, al testo tradotto. Ciò detto, è chiaro che tale tempo è da considerarsi in stretta relazione con la natura e il numero delle correzioni, le cui varietà possono influenzare pesantemente questo dato.

Tempo necessario per la traduzione Suddetto criterio è obiettivamente di scarsa utilità alla valutazione della qualità traduttiva, dal momento che deve prendere in considerazione tutta la tempistica necessaria, dal momento della ricezione dell'ordine alla sua consegna, valutando pertanto diversi fattori quali l'organizzazione della traduzione, della revisione e del post-editing, flusso di lavoro dell'agenzia che prende in carico il progetto, performance della strumentazione utilizzata e così via. Di conseguenza, risulta quantomeno fuorviante considerare questi dati come indicativi della qualità traduttiva, a meno che non si conduca un'indagine di mercato atta a valutare determinati standard produttivi, caso in cui costituirebbero invece un elemento di fondamentale importanza.

Livello linguistico Tale criterio di valutazione, come il nome suggerisce, prevede l'analisi di fenomeni prettamente linguistici, quali l'equivalenza semantica, la corretta traduzione di sintagmi nominali, la corretta identificazione e traduzione degli argomenti del verbo, corretto utilizzo di articoli e punteggiatura e corretta identificazione dei referenti testuali, procedendo quindi all'analisi di eventuali errori e regole grammaticali della lingua target non rispettate.

Livello operativo Certamente, valutare la qualità traduttiva di un sistema MT attraverso la sua esecuzione e analisi degli eventuali requisiti che ci si aspetta siano soddisfatti è un valido metodo di valutazione, ma piuttosto vago poiché sottoposto a esigenze e necessità di volta in volta diverse, senza contare gli ingenti costi che da esso possono derivare.

Se da un lato quindi la suddetta macrovalutazione proposta si prefigge l'obiettivo di valutare la qualità traduttiva in sé, senza indagare nelle cause o nelle possibili soluzioni, dall'altro lato, in direzione opposta, sempre Van Slype (ivi: 116) propone anche una microvalutazione, il cui fine ultimo è quello di quantificare gli errori, determinarne le cause e infine adoperarsi per trovare possibili rimedi e soluzioni atte a migliorare le performance traduttive. Tale valutazione si snoda lungo cinque livelli, descritti a seguire.

Livello grammaticale Consta della frequenza di errori corretti, tramite post-editing, secondo la tipologia di errore (morfologico, sintattico, semantico, ecc.). Sebbene questo sembri il metodo più naturalmente ovvio di affrontare eventuali carenze in termini di qualità produttiva, nella pratica si rivela essere inadeguato poiché non apporta alcun tipo di miglioramento alle prestazioni di un sistema MT (ivi: 173).

Livello formale Equivale alla valutazione degli errori in base al tipo di correzione effettuata da un post-editor, ovvero aggiunta o eliminazione di una o più parole, riposizionamento delle stesse e così via. Sebbene questo metodo permetta di valutare in maniera qualitativa e quantitativa il lavoro effettivo di un post-editor, si rivela a sua volta inefficace, dal momento che i suoi risultati sono facilmente influenzabili dall'atteggiamento che diversi post-editor possono dimostrare nei confronti della machine translation, in virtù del quale alcuni faranno del loro meglio per limitare al minimo il numero di correzioni apportabili, mentre altri opereranno per un numero di correzioni maggiori e non necessario (*ibid.*).

Livello causale Questo metodo di valutazioni prevede l'analisi delle cause delle performance carenti di un sistema MT. Per quanto suddetto metodo indaghi in maniera più approfondita nei punti deboli di un sistema MT, anche questo metodo si rivela inefficace, dal momento che il tipo di analisi condotta è comunque superficiale e limitata a un'analisi del testo source e i risultati da esso riportati non si dimostrano utili a fornire la base per ulteriori studi (ivi: 174).

Margine di miglioramento A questo livello corrisponde l'effettivo scopo della microvalutazione, ovvero l'analisi di differenti soluzioni apportabili al sistema MT, al fine di trovare quella più adatta a prevenire il ripetersi di determinati errori costanti, nonché un'analisi delle risorse necessarie a implementare suddette soluzioni, stimando quindi i costi necessari. Dei cinque livelli, questo si dimostra essere quello più adeguato per una strategia di miglioramento delle prestazioni di un sistema MT, in virtù della sua analisi di costi e risultati ottenibili da ogni soluzione (*ibid.*).

Implementazione delle migliorie Sebbene lo scopo della microvalutazione sia quello di individuare le cause di performance carenti di un sistema MT e trovare le relative soluzioni, quella dell'implementazione vera e propria di suddette soluzioni

sembra una fase quantomeno al di là dei reali obiettivi di una valutazione e per tale motivo se ne sconsiglia l'attuazione, per lo meno entro i limiti di questo studio, come suggerisce il succitato Van Slype (ivi: 175).

Sebbene quello della qualità traduttiva sia un parametro di fondamentale importanza, la *conditio sine qua non* concernente l'utilizzo di un determinato sistema MT piuttosto che un altro, è doveroso puntualizzare quanto esso non sia sufficiente a determinare l'effettiva utilità della machine translation, dal momento che tale discorso richiede un approccio che parta da più prospettive. Ad esempio, se da un lato, gli sviluppatori di suddetti sistemi sono interessati alla corretta resa di strutture sintattiche e semantiche, i clienti che scelgono di avvalersi di sistemi MT vorranno, dal canto loro, la sicurezza che tali strumenti apportino un effettivo risparmio di denaro. Dal canto suo, invece, il traduttore avrà un maggiore interesse nel verificare che machine translation e post-editing comportino un aumento della sua produttività, a fronte di tempi di traduzione minori necessari a compensare tariffe traduttive più basse. Infine, c'è la prospettiva dell'utente finale: per quest'ultimo, la traduzione adatta è quella che soddisfa i criteri di utilizzabilità, adempiendo allo scopo ultimo del testo con il massimo effetto comunicativo. Posto pertanto che l'utilizzatore ultimo sia, per esempio, un utente medio, per lui determinati testi "usa e getta" non necessiteranno dei più alti standard di qualità, poiché l'utilità di quel documento si esaurisce nel momento in cui viene letto, ragion per cui i suoi standard di consegna e qualità del testo tradotto differiranno enormemente da quelli di un Language Service Provider.

In base ai punti di vista valutati, è possibile quindi distinguere tra approcci formali e approcci funzionali (Vasconcellos, 1988). Dei due, come è lecito aspettarsi, la seconda tipologia è quella che restituisce i risultati più rilevanti. Gli approcci formali, di cui l'esempio più famoso è il succitato rapporto ALPAC, risultano limitati nel loro essere poiché si concentrano prevalentemente su uno o due aspetti per valutazione e rivolgono la loro attenzione al solo output della machine translation, definito *raw*, ovvero esente da alcun tipo di lavoro di revisione o post-editing, ignorando quindi il prezioso intervento del traduttore post-editor, il cui lavoro è sì utile al testo tradotto stesso ma, in maniera più ampia, al potenziamento e sviluppo degli stessi sistemi

MT. Il pregio degli approcci funzionali è invece quello di indagare sull'effettiva utilità di suddetti sistemi, limitatamente allo scopo per cui sono stati creati. Per ottenere risultati di questo tipo è necessario rivolgersi all'utente e fruitore della machine translation, poiché è solo a lui che spetta determinare le potenzialità di suddetti sistemi e in che misura questi costituiscono un asset nella propria professione. Al di là di tale indagine, poi, una visione d'insieme più ampia necessita dell'analisi di ulteriori aspetti, quali i dizionari utilizzati dal sistema, l'ambiente di lavoro, la figura stessa del traduttore che si avvale della machine translation e il supporto degli sviluppatori (ivi: 210).

Analizzare la struttura delle risorse linguistiche a disposizione consente di avere un'idea delle potenzialità del sistema, come ad esempio la possibilità di utilizzare dizionari settoriali o sotto-dizionari definibili dall'utenza, così come l'implementazione di informazioni di carattere semantico e sintattico indispensabili per poter operare la giusta scelta in base al contesto comunicativo. Per quanto concerne l'ambiente lavorativo, è importante valutare l'efficacia degli strumenti messi a disposizione del post-editor, il quale deve essere considerato come l'anello di congiunzione di una catena di lavorazione più ampia, nella quale il passaggio di informazioni e materiali e da un stadio all'altro del processo traduttivo avviene nel modo più fluido possibile. Come detto poc'anzi, limitatamente al ruolo del traduttore, il suo contributo è di fondamentale importanza poiché l'utilizzo giornaliero dei sistemi MT consente a questi di migliorare le prestazioni tramite training continuo, aggiornamento dei dizionari e potenziamento dell'algoritmo, consentendo altresì l'individuazione di errori o malfunzionamenti ricorrenti. Infine, è senz'altro produttivo valutare il supporto degli sviluppatori, dal momento che un sistema MT allo stato dell'arte è un sistema MT costantemente aggiornato e mantenuto funzionante.

Pertanto, avvalendosi dei criteri elencati in precedenza e includendo le considerazioni esposte poc'anzi, è possibile ricavare dati utili e interessanti per una valutazione obiettiva ed esaustiva dell'output di un sistema MT, sia che si vogliano testare le potenzialità di tale strumento, sia che si vogliano valutare prestazioni del post-editing ed effetti sugli standard di produttività. A tale proposito, è opportuno analizzare i risultati di test empirici che si muovono in tale direzione, ovvero

confrontando standard di qualità e produttività tra testi tradotti da traduttori professionisti, e quindi senza l'ausilio della machine translation, e testi machine translated e poi sottoposti a postediting (Fiederer e O'Brien, 2009). Per una valutazione di questo tipo, su frasi estratte dal manuale utente di un software, i criteri adottati sono stati quelli di intelligibilità/chiarezza, fedeltà/accuratezza e stile. Di questi tre, sicuramente lo stile è il parametro che genera quesiti e risultati tra i più interessanti. Per valutazione dello stile si intende valutare la qualità traduttiva in funzione dell'appropriatezza al contesto, nonché allo *skopos*, del linguaggio utilizzato. Sebbene, dei tre, suddetto criterio sia quello soggetto a un maggiore grado di soggettività, inserirlo in questa valutazione risulta del tutto coerente con l'idea che, per quanto a un post-editor venga richiesto di alterare il meno possibile l'output della machine translation, qualora questo sia accettabile, con il solo scopo di innalzarne lo stile, l'istinto è comunque quello di revisionare l'output della MT per renderlo equivalente a quello di una traduzione prodotta senza machine translation.

Fatte queste premesse, come puntualizza O'Brien (ivi), i risultati predicibili dovrebbero essere in netto favore per i traduttori per tutti e tre i parametri, restituendo invece punteggi inferiori per i post-editor. Al contrario, come si evince in figura 2.1, i risultati riportati da questo studio erano in netta contraddizione con le aspettative, almeno per quanto concerneva i primi due parametri.

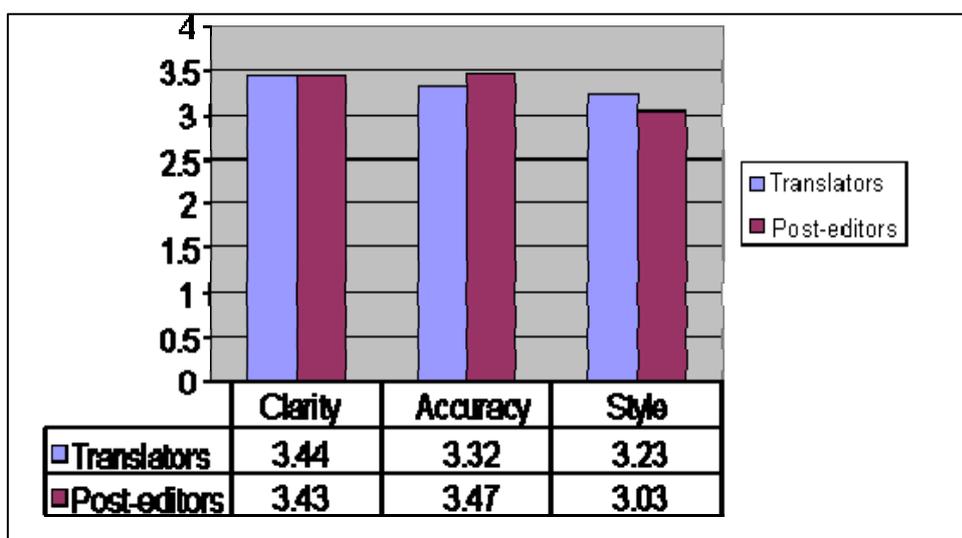


Figura 2.1 Punteggi medi per traduttori e post-editor nei tre criteri adottati per la valutazione. (O'Brien, 2009)

Se da un lato i punteggi di traduttori e post-editor sono pressoché equivalenti per quanto riguarda la chiarezza, dall'altro, la performance dei post-editor è stata giudicata migliore per quanto concerne l'accuratezza. Ciò è probabilmente dovuto al fatto che, con ogni probabilità, i post-editor erano più intenti a riportare in maniera esatta ed equivalente tutte le informazioni del testo source, al contrario dei traduttori, per cui è presumibile che la preoccupazione maggiore fosse per la resa stilistica. Tale considerazione trova la sua conferma nei punteggi attribuiti a traduttori e post-editor sullo stile: il punteggio medio inferiore assegnato ai post-editor è presumibilmente dovuto al fatto che per un post-editor, a patto che l'output della machine translation si conformi alle regole della lingua target, i cambiamenti da apportare devono essere ridotti allo stretto necessario, così comportando, inevitabilmente, una resa stilistica che in alcuni punti risulta imitativa dello stile della lingua source. Un traduttore, al contrario, farà tutto il possibile perché il suo prodotto venga letto come un testo indipendente e non una traduzione. A conclusione di quanto esposto tramite i risultati del suddetto esperimento qui riportato, il lettore è invitato a ricordare tali osservazioni, poiché saranno oggetto di ulteriori considerazioni a supporto di quanto si mostrerà nel capitolo successivo.

2.2 POST-EDITING E INCREMENTO DEGLI STANDARD DI PRODUZIONE: METODOLOGIE DI STUDIO DELLA CORRELAZIONE TRA QUALITÀ E PRODUTTIVITÀ

Quanto detto finora ha fornito dati certi a riprova del fatto che l'utilizzo congiunto di traduzione automatica e post-editing apporta vantaggi non indifferenti per i traduttori che scelgono di avvalersene, in maniera più o meno equivalente a seconda dell'aspetto sul quale si decide di concentrarsi, ma generalmente accettabili o quantomeno tali da giustificare l'impiego di suddetta tecnologia. Per potere, tuttavia, descrivere in maniera ancor più dettagliata i vantaggi che il post-editing consente di ottenere, pare opportuno concentrarsi su ulteriori aspetti, i quali, oltre alla comprovata accuratezza, sono di fondamentale importanza nella valutazione dell'efficacia traduttiva, poiché consentono di migliorare la qualità traduttiva e, al tempo stesso, ridurre i costi di produzione, che poi sono, tra l'altro, le motivazioni principali addotte dagli sviluppatori di sistemi MT. Tali aspetti sono un minore

quantitativo di tempo richiesto in, generalmente, tutte le fasi del processo traduttivo, con tempi di produzione conseguentemente minori, e uno sforzo cognitivo da parte del traduttore presumibilmente minore.

A tal proposito, sarà di sicura utilità ai fini di questa indagine fornire una definizione e descrizione delle differenti fasi di cui consta il processo traduttivo, come queste figurano in Carl, Kay e Jensen (2010), ivi denominate *gisting*, *drafting* e *post-editing*, alla luce delle differenze sostanziali in termini di attività cognitiva e meccanica osservabili nelle differenti fasi, attraverso le quali sarà poi possibile comprendere in maniera estensiva gli studi e indagini analitiche che si presenteranno a seguire.

Le performance di un traduttore variano notevolmente in funzione di parametri generalmente soggettivi, ma, in maniera piuttosto chiara e definita, è possibile isolare le tre fasi di cui consta il processo traduttivo, ovvero le succitate fasi di *gisting*, *drafting* e *post-editing*. La prima fase consta della lettura, acquisizione e comprensione del testo source e quindi lo sforzo cognitivo richiesto al traduttore sarà proporzionale al grado di specializzazione del testo, allo stile e alla terminologia, nonché, presumibilmente, al suo grado di esperienza. La fase di *drafting* consiste nella stesura della traduzione vera e propria, sotto forma di bozza da rivedere e correggere. In questa fase sarà possibile osservare, come si vedrà a seguire, un tipo di attività marcatamente differente da quelle osservabili nelle altre due fasi. La terza fase, ovvero il *post-editing*, consiste nella rilettura, nella correzione di eventuali refusi o frasi intere da riorganizzare e ristrutturare su diversi piani linguistici, in virtù di una maggiore comprensione da parte del traduttore arrivato a questo stadio (ivi).

Per comprendere meglio i fenomeni che si intende analizzare, è possibile prendere nota di quanto rappresentato nella figura a seguire:

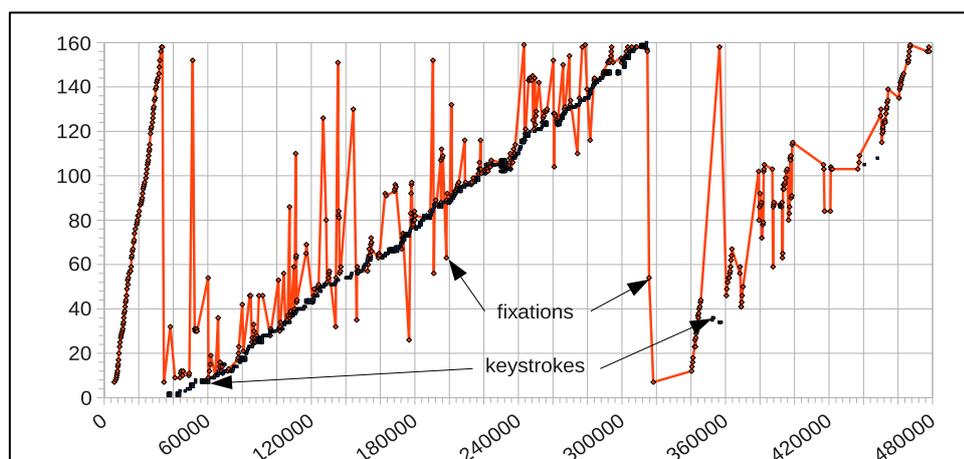


Figura 2.2 Grafico del processo traduttivo e registrazione del movimento degli occhi e della pressione dei tasti sulla tastiera di un computer. L'asse orizzontale rappresenta la durata di tempo in millisecondi necessaria impiegata per tradurre, mentre l'asse verticale rappresenta, dal basso verso l'alto, il numero di parole del testo source. La pressione dei tasti registrata sulla tastiera (*keystrokes*) è rappresentata da un punto nero mentre la linea rossa descrive il movimento degli occhi sul testo source e i punti lungo suddetta linea equivalgono alle fissazioni oculari (*fixations*), ovvero allo sguardo fisso prolungato del traduttore su un determinato punto del testo source. (Carl, Kay e Jensen, 2010)

Nel grafico in figura, nel quale sono state registrate le attività di un traduttore a lavoro su una traduzione e il suo progresso nella stessa, sono chiaramente distinguibili le tre fasi summenzionate poc'anzi. I primi quaranta secondi registrati sono stati impiegati dal traduttore per la sola fase di comprensione e assimilazione del testo source e tale dato è rafforzato dal numero di volte in cui, nell'arco di tempo preso ora in esame, lo sguardo fisso del traduttore è ricaduto su una parola precisa del testo, che pare sia stato letto dall'inizio alla fine. Successivamente, si assiste a un'intensa attività di digitazione sulla tastiera, indicata dal numero di pressioni registrate dal secondo 40 al secondo 320 circa: tale porzione di tempo è quella dedicata alla fase di *drafting*. Anche il numero di spostamenti dello sguardo nella porzione di testo osservata dal traduttore conferma questo dato poiché è indice di movimenti dell'occhio che vanno dal testo source attualmente tradotto alla finestra in cui è probabile che si stia digitando il testo target. Infine, dal secondo 320 al secondo 480, è possibile prendere nota di quanto accade nella fase di *post-editing*, in cui il traduttore rilegge gran parte del testo e solo alcune sporadiche pressioni dei tasti vengono registrate.

Questo tipo di dati permette di analizzare i livelli di produttività traduttiva da prospettive diverse e considerevolmente innovative, dal momento che possono

fornire informazioni interessanti circa sforzo cognitivo e incremento della qualità finale del prodotto, avvalorando l'ipotesi secondo cui il post-editing apporta dei reali vantaggi al traduttore. A tal fine, è bene indagare in modo più approfondito l'analisi dei dati fornitici dal monitoraggio dei summenzionati *keystrokes* e *fixations*, avvalendosi del contributo di preziosi studi circa queste attività.

2.2.1 POST-EDITING E INDICATORI DI INCREMENTO PRODUTTIVO: ANALISI DELLE PRESTAZIONI IN FUNZIONE DELL'INPUT SU MOUSE E TASTIERA

Il post-editing richiede, come detto svariate volte in questa sede, la commistione di diverse attività, cerebrali e meccaniche, atte a bilanciare sforzo cognitivo necessario alla comprensione del testo, valutazione dell'output della machine translation e generazione del testo target definitivo. Tali attività si riflettono in una reale attività fisica su tastiera, nonché mouse, descrivibili qualitativamente e quantitativamente a seconda della fase del processo traduttivo analizzato. Green, Heer e Manning (2013) rintracciano nel loro studio trend e comportamenti ricorrenti riconducibili a differenti livelli di attività cerebrale e meccanica. Nella figura a pagina seguente, ad esempio, vengono presi in considerazione i movimenti del mouse sopra porzioni di un testo inglese preso in esame quale metro di valutazione dell'attività traduttiva di un traduttore arabo, di un traduttore francese e di un traduttore tedesco.

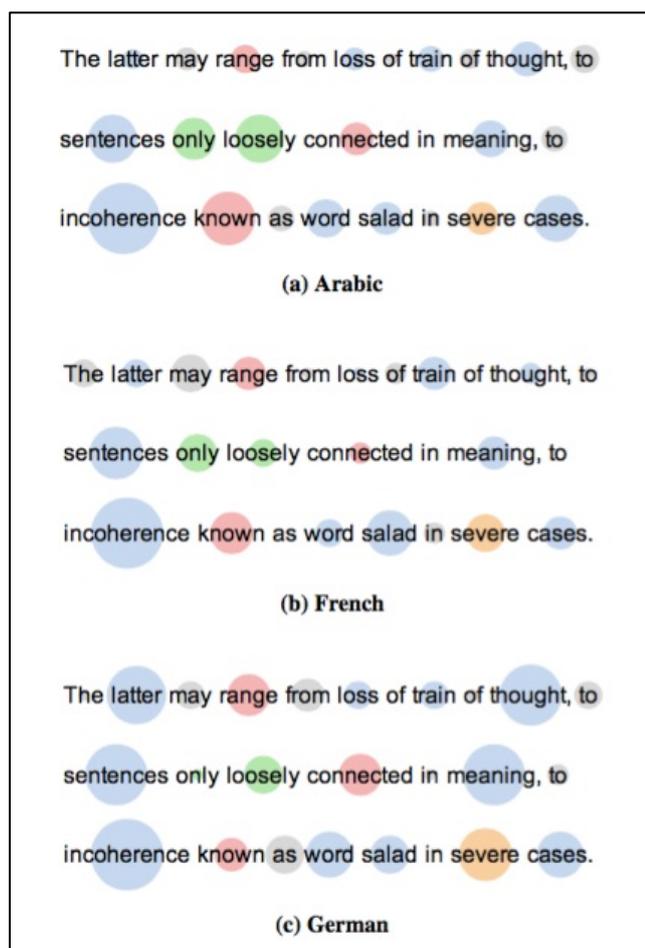


Figura 2.3 Frequenza dei movimenti del mouse su porzione di testo inglese per le tre lingue prese in esame da Green, Heer e Manning (2013). La frequenza è indicata dall'ampiezza dell'area colorata, dove differenti colori rappresentano differenti categorie grammaticali: in blu sono indicati i sostantivi, in rosso sono indicati i verbi, in arancione gli aggettivi, in verde gli avverbi e in grigio altre categorie grammaticali, quali ad esempio articoli o preposizioni.

Come emerge dalla figura di cui sopra, in tutte e tre le lingue oggetto dello studio, i nomi prevalevano quali nodo concettuale importante, seguiti da avverbi e, in minor misura, da verbi. Il fatto che i movimenti del mouse si siano concentrati prevalentemente su queste aree in tutte e tre le lingue suggerisce una probabile attività cognitiva e assimilativa del testo source che può avere effetti più o meno sensibili sullo svolgimento della traduzione, con una persistenza significativa tra le tre lingue tale da suggerire un trend ricorrente (ivi).

Disponendo dei dati di cui sopra, è inoltre possibile valutare l'attività del traduttore su mouse e tastiera ponendo a confronto il lavoro di un traduttore che non

ha usufruito di machine translation con uno che ne ha fatto uso, come rappresentato dalla seguente figura, tratta sempre dal medesimo studio del 2013.

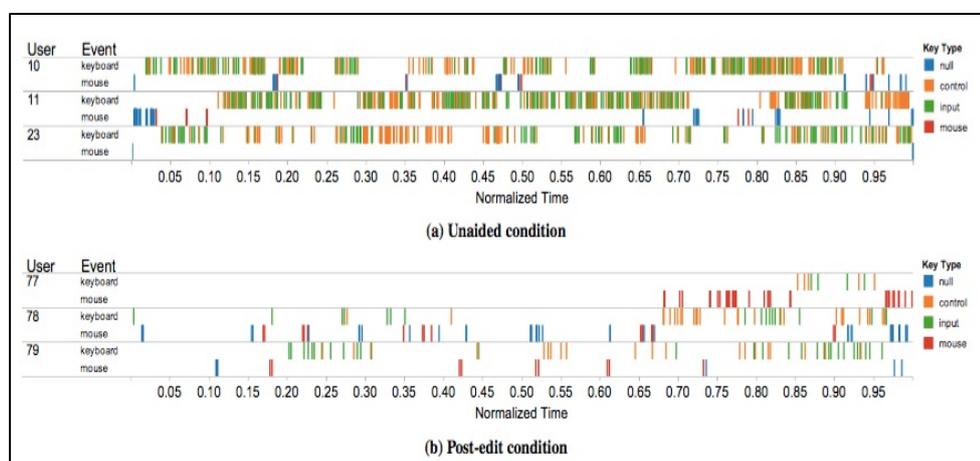


Figura 2.4 Registrazione dell'attività su mouse e tastiera della traduzione del testo source. In figura l'attività su mouse e tastiera della traduzione del testo mostrato in Figura 2.3 tramite traduzione (a) e tramite post-editing (b). La tipologia di attività su tastiera è stata differenziata in base al colore: i tasti di scorrimento (freccia in su/giù) sono evidenziati in arancione; l'inserimento di testo è evidenziato in verde; l'attività registrata sui pulsanti del mouse è evidenziata in rosso. I casi in cui il programma non è riuscito a comprendere la tipologia di attività eseguita sono evidenziati in blu (Carl, Heer e Manning, 2013).

Dall'analisi della registrazione svolta da Carl, Heer e Manning sull'attività su mouse e tastiera emergono due andamenti nettamente contraddistinti tra le due situazioni. Se da un lato nel primo caso è possibile differenziare in maniera abbastanza intuitiva tra le tre fasi del processo traduttivo, dove ad esempio pause iniziali e intensa attività su tastiera sono caratteristiche delle fasi di *gisting* e *drafting*, nel secondo caso il confine tra le tre fasi si fa piuttosto labile e le pause sono maggiori. Ciò è dovuto al fatto che, con ogni probabilità, il traduttore ha scelto di non scartare definitivamente l'output della machine translation e di conseguenza l'attività su tastiera risulta minore, avanzando l'idea che il post-editing sia un'attività più passiva rispetto alla traduzione (ivi). Inoltre, un maggiore utilizzo del mouse e dei tasti di scorrimento permette di avanzare l'ipotesi che il post-editor si sposti con considerevole agilità lungo il testo per andare a correggere i soli punti che necessitano il suo intervento, con un conseguente risparmio di tempo non indifferente.

Di conseguenza, un fattore che gioca un ruolo importante nell'analisi dell'attività di digitazione su una tastiera di un traduttore a lavoro non è solo la frequenza della

singola pressione esercitata sul singolo tasto in sé, quanto le pause che intercorrono tra una pressione e l'altra, postulando il fatto che le pause siano effettivi indicatori di un quantificabile sforzo cognitivo durante la fase di post-editing. Sebbene gran parte degli studi condotti sulla correlazione tra pause e sforzo cognitivo sembrino fallire nel confermare quanto appena affermato, sostenendo, quantomeno, che i dati a portata di mano non sono sufficienti a supportare questa tesi, è tuttavia opportuno riportare alcune delle nozioni ivi affrontate.

In Séguinot (1989: 31), un'attività di digitazione insolitamente lenta è definibile come esitazione mentre per pausa vera e propria si deve intendere un'interruzione nella digitazione e sempre la studiosa trova che una maggiore concentrazione di quest'ultime occorra in corrispondenza della fine di una frase o di un paragrafo, tra due proposizioni indipendenti, prima o dopo proposizioni subordinate, prima di sintagmi, prima di soggetto e predicato, alla fine di una riga e prima o durante la lettura di una parola. Da un punto di vista quantitativo, invece, se una pausa di circa 0,20 è da considerarsi coerente con la velocità media di digitazione di diversi individui, una pausa che ammonta a un secondo è da considerarsi come un vero e proprio ritardo nella velocità di produzione testuale (Jakobsen, 1998 cit. in O' Brien, 2006: 3).

Sebbene queste definizioni siano coerenti con quanto riscontrato in diversi studi sull'argomento, risulta tuttavia difficile estrapolare dati affidabili dall'analisi delle pause durante l'attività di post-editing, anche qualora si confrontino testi sottoposti a pratiche utili a incrementare la produttività quali l'implementazione delle regole di *controlled language* (CL), di cui si discuterà più avanti. Ciò è dovuto al fatto che le differenze quantitative e qualitative riscontrabili nelle performance di diversi individui non sono tali da poter delineare un trend ricorrente, alla luce di variazioni non correlabili tra loro, tanto tra soggetti diversi quanto tra performance dello stesso soggetto, nonché tra un segmento e l'altro e nell'ambito dello stesso segmento (ivi: 11). Inoltre, diversi studi sollevano ragionevoli dubbi circa i fattori determinanti l'occorrenza di una pausa, i quali possono essere di natura fisiologica, ovvero l'effettiva necessità del traduttore di concedersi una pausa, nonché organizzativa, ovvero la pianificazione della resa del contenuto successivo a quello già tradotto (ivi: 16). Del resto, in diverse occasioni O'Brien riscontra anche un frequente utilizzo dei

cursori di movimento per scorrere da una parola all'altra del testo: potrebbe questa attività essere ricondotta a un meccanismo sostitutivo alla pausa che consente di scansionare e comprendere il testo parola per parola, analogamente allo scorrere della penna su una pagina durante la lettura? Tale prospettiva risulta tanto affascinante quanto plausibile, cionondimeno complicando l'interpretazione delle interruzioni nel ritmo produttivo.

Tali inconcludenze sono una riprova del fatto che, per quanto sia effettivamente possibile postulare un legame tra pause e sforzo cognitivo, gli studi attualmente in possesso non sono sufficienti a sostenere tale tesi, per la quale si auspicano invece studi più approfonditi che tuttavia esulano dallo scopo dello studio presente. Pertanto, riconoscendo comunque un limite concreto nella quantificazione dello sforzo cognitivo durante il post-editing valutando i fenomeni finora descritti, l'analisi dell'attività su tastiera e mouse consente per lo meno di riscontrare un guadagno in termini di tempo e digitazione su tastiera, a fronte di una qualità dell'output della MT soddisfacente.

2.2.2 CORRELARE QUALITÀ E SFORZO COGNITIVO NEL POST-EDITING: EYETRACKING E FISSAZIONI OCULARI

A sostenere la tesi secondo cui l'analisi del movimento dell'occhio e il suo soffermarsi su un determinato punto permetta di ricavare dati utili circa lo sforzo cognitivo, vi è la concezione che, in tali momenti di osservazione prolungata, scientificamente denominati fissazioni oculari, il cervello sia impegnato in un'intensa attività cognitiva di decodifica della frase in quel momento osservata. La letteratura a riguardo è piuttosto vasta, soprattutto in correlazione agli studi sulla velocità di lettura. Rapportando queste considerazioni alla valutazione della qualità dell'output della machine translation e del post-editing, oggetto di molti studi è l'indagine circa un'ipotetica correlazione tra movimento oculare e qualità dell'output della MT.

Alla base di molti studi in materia sottostà l'idea che quanto più alto è il livello di qualità dell'output della machine translation tanto maggiore sarà la facilità con cui è possibile leggere suddetta traduzione. D'altronde, guardando la cosa da una prospettiva opposta, è altrettanto vero che la facilità con cui è possibile leggere e

comprendere l'output della machine translation influenza positivamente, o negativamente in alcuni casi, il giudizio qualitativo che si ha di essa, dando vita a quello che si può definire un circolo virtuoso. Chiaramente, le conoscenze pregresse di cui un'eventuale traduttore dispone in materia influenzeranno tali valutazioni, dal momento che, postulando l'inesistenza di alcuna conoscenza personale riguardo il contenuto del testo tradotto, in presenza di un testo malamente tradotto tramite machine translation, lo sforzo cognitivo necessario alla comprensione del testo sarà inversamente proporzionale alla sua qualità.

Studi empirici di questo tipo hanno lo scopo di correlare quanto detto sopra con movimento oculare e durata media delle fissazioni su un dato passo del testo e sul testo nella sua interezza. A tal proposito, al fine di avvalorare la tesi summenzionata all'inizio del paragrafo, si presentano i risultati di uno studio empirico condotto da Doherty, O'Brien e Carl (2010) la cui rilevanza, nonché metodologia adottata, si ritiene essere coerente e utile ai fini di questo studio. Uno dei primi dati presi in considerazione è la durata media dedicata da un individuo all'osservazione di determinati punti di interesse su schermo, ovvero, nella fattispecie, ogni singola frase. Prevedibilmente, i risultati mostrano che, effettivamente, il tempo dedicato alla lettura dell'output della MT di qualità inferiore risulta essere maggiore, come riportato in figura 2.5.

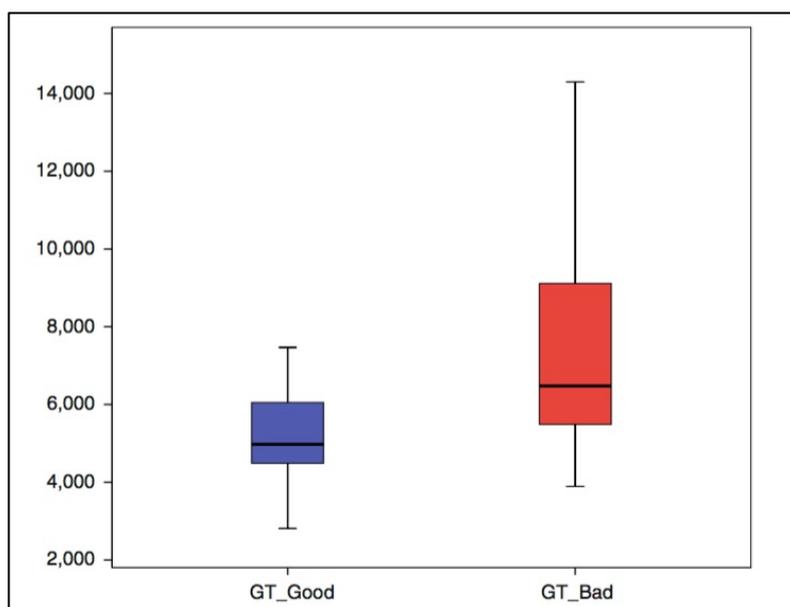


Figura 2.5 Durata media (GT) dedicata all'osservazione di ogni frase dei due differenti output della machine translation. In blu (GT_Good) è indicato il dato relativo alla traduzione di alta qualità; in rosso

(GT_Bad) è indicato il dato relativo alla traduzione di scarsa qualità (Doherty, O'Brien e Carl, 2010: 6).

Analogamente, il conteggio delle volte in cui lo sguardo dell'individuo si è fissato su un'area d'interesse restituisce valori concordi a quelli precedenti, dove, pertanto, al testo malamente tradotto corrispondono un numero maggiore di fissazioni oculari come mostrato in figura 2.6.

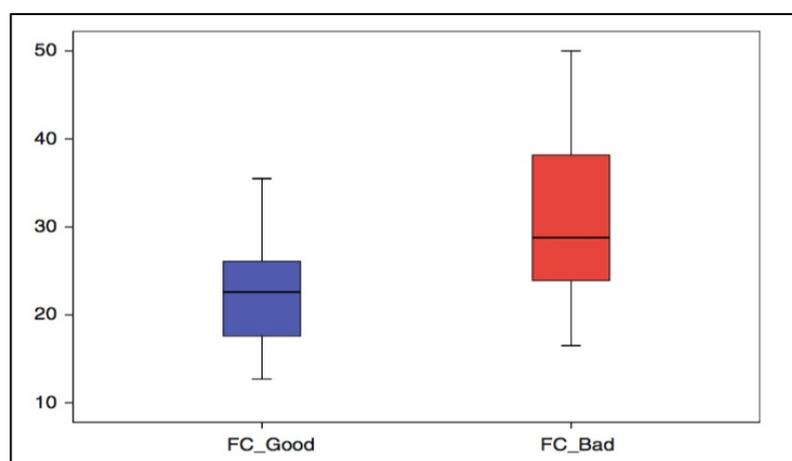


Figura 2.6 Conteggio medio delle fissazioni oculari. (Doherty, O'Brien e Carl, 2010: 7).

Un dato estremamente interessante, che emerge da studi di questa natura, è l'ottima correlazione che le misurazioni del movimento oculare sembrano avere con diversi parametri di valutazione automatica dell'output della machine translation, in particolar modo con TER. È questo il caso di uno studio condotto nel 2011, nel quale utilizzando un eye tracker è stato possibile calcolare numero e durata media delle fissazioni durante la fase di post-editing, ponendo poi questi dati in relazione con le valutazioni condotte tramite TER (O'Brien, 2011).

Come prevedibile, il punto d'interesse su cui è stato registrato il numero maggiore di fissazioni è la porzione di schermo su cui figuravano testo source e testo target, come mostrato in figura 2.7.

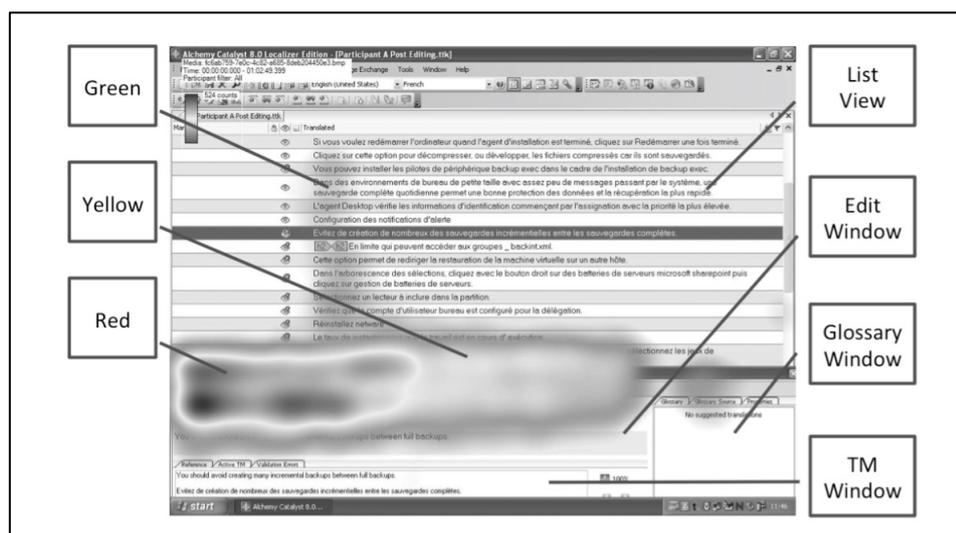


Figura 2.7 Interfaccia utente del CAT tool Catalyst, utilizzato dai partecipanti allo studio per svolgere la traduzione assegnata. L'area in rosso indica il punto d'interesse in cui è stato registrato il maggior numero di fissazioni, ovvero la finestra *Edit*. A seguire le aree oggetto del maggior numero di fissazioni sono la vista *List* in cui sono elencati i segmenti target e il loro status e la finestra *Glossary* (O'Brien, 2010: 11).

A sostenere la validità del succitato studio vi è la tesi secondo cui maggiore è la durata delle fissazioni, maggiore è lo sforzo cognitivo. Analogamente, maggiore è il numero di fissazioni, maggiore è lo sforzo cognitivo in atto nel corso del processo (ivi: 15-17). Effettivamente, come riportano i risultati dello studio di O'Brien, a segmenti valutati negativamente attraverso l'utilizzo di parametri di valutazione automatica corrisponde una durata maggiore, espressa in secondi, delle fissazioni, così come un numero maggiore delle stesse.

Ciò detto, appare ragionevole concludere che la misurazione del movimento oculare quale indicatore dello sforzo cognitivo individuale durante il post-editing restituisce risultati validi a sostanziare la tesi di cui sopra, secondo cui a una qualità maggiore dell'output della traduzione automatica corrisponde un minor sforzo cognitivo, alla luce inoltre della correlazione con diversi parametri di valutazione automatica dell'output della machine translation.

2.3 POTENZIAMENTO E OTTIMIZZAZIONE DELLE PRESTAZIONI: PRE-EDITING, POST-EDITING AUTOMATIZZATO E CONFIDENCE SCORES

Le variabili che possono incidere negativamente sulle prestazioni di un motore di traduzione automatica e di conseguenza sul lavoro seguente che il post-editor dovrà compiere sull'output della MT sono molte. Per ovviare a tali scenari, i quali rischiano di minare l'efficacia della MT e del post-editing, una parte piuttosto attiva della letteratura sulla traduzione automatica si concentra sull'elaborazione di meccanismi e strategie in grado di potenziare e ottimizzare le prestazioni di suddetti sistemi e dei post-editor stessi, automatizzando i task richiesti. Si sta qui parlando dell'impiego di set di regole noti come *Controlled Language* e dell'implementazione del post-editing automatizzato e dei *confidence score*.

2.3.1 PRE-EDITING E OTTIMIZZAZIONE DEI TESTI ATTRAVERSO L'IMPLEMENTAZIONE DI SET DI REGOLE CONTROLLED LANGUAGE (CL)

Produrre e ottimizzare i testi perché aderiscano a determinate linee guida è una componente fondamentale della produzione testuale di società, istituzioni e aziende, la cui necessità è ormai data come imprescindibile, poiché in essa si riflette l'identità stessa dell'autore, ovvero della società per cui è stato scritto quel determinato testo. A tal proposito si può quindi parlare di guide di stile nelle quali ogni società fissa i propri requisiti e norme redazionali, tanto generali quanto specifici. Da un punto di vista linguistico, si parla in questo caso dell'impiego di un linguaggio standard, un uso coerente e adeguato della terminologia, nonché il divieto di utilizzare espressioni idiomatiche, ovvero tutta una serie di accorgimenti atti ad accrescere la comprensibilità di un testo e ridurre al minimo l'ambiguità.

Questo prescrittivismismo si può estendere anche alla traduzione automatica, al fine di creare dei testi in grado di essere facilmente analizzabili e traducibili da un sistema MT. Analogamente alle guide di stile e già per la traduzione assistita, nel corso degli anni, in seguito a diversi studi, sono state elaborati set di regole che insieme danno origine a un linguaggio definito *Controlled language (CL)*, ovvero un linguaggio imitativo di quelli naturali, la cui grammatica e i dizionari hanno subito processi di controllo e restrizione al fine di ridurre la complessità e ambiguità che invece caratterizza tutte le lingue naturali (Gough e Way, 2003). I differenti set CL vanno ad

agire su differenti aspetti linguistici, come per esempio l'utilizzo della *-ing clauses* o delle proposizioni relative, come mostrato nei seguenti esempi (Aikawa *at al.*, 2007):

(1) Tolerance limits are developed with environment owners *before allowing* each new environment to access the network.

Tolerance limits are developed with environment owners *before each new environment is allowed* to access the network.

(2) Use only fonts *optimized* for display on the Web.

Use only fonts *that are optimized* for display on the Web.

Sebbene, alla luce della reversibilità delle lingue in memorie e motori di traduzione, questo possa rivelarsi un aspetto negativo della CL, un'operazione di questo tipo è per forza di cose applicabile e limitabile al solo testo source, dal momento che, stando a quanto riportato nella letteratura specifica, finora nessuna coppia di segmenti source-target è in grado di conformarsi alle regole di almeno un set di Controlled Language. Inoltre, qualora si volesse comunque tentare di trovare coppie simili, alla luce dell'apparente impossibilità, questo diverrebbe un compito piuttosto arduo, per il quale non sussistono garanzie della sua riuscita e dell'effettivo incremento della produzione (Gough e Way. 2004). Del resto, determinate regole valide per una lingua possono non esserlo per un'altra, motivo per cui risulta impossibile trovare un set di CL restituente risultati ottimali su diverse lingue. (2008: "An introduction to the use of controlled language"). Pertanto, come dimostrano gli studi, applicando tali set all'input della machine translation, aumentandone la traducibilità, è possibile ottenere un output della MT di maggiore qualità e, di conseguenza, maggiore produttività durante il post-editing, in quello che può essere definito un circolo virtuoso, come mostrato in figura 2.8.

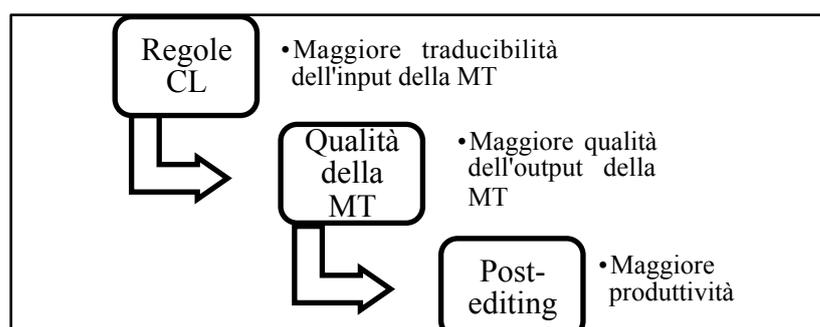


Figura 2.8 Regole Controlled Language (CL), qualità della MT e produttività del post-editing. (Aikawa *et al.*, 2007)

D'altro canto, è tuttavia opportuno riportare gli eventuali effetti controproducenti derivanti dall'utilizzo della CL, ovvero il fatto che l'imposizione di regole ferree e poche scelte traduttive disponibili possa essere percepita come un limite all'attività del traduttore, nonché la necessità di uniformare ai set CL prescelti tutto l'eventuale materiale di riferimento di cui constano i database di una società (2012: "Controlled language").

Ciononostante, alla luce del comunque innegabile apporto della CL alla traduzione automatica, come ragionevole conseguenza, per quelle compagnie che adottano set di regole CL diviene di fondamentale importanza dotarsi di applicazioni specifiche utili alla verifica dell'effettiva conformazione dei testi prodotti ai set di regole adottati. Tali applicazioni sono equiparabili ai ben noti correttori di cui sono dotati i più comuni programmi di scrittura disponibili su mercato e constano di una componente che individua gli errori e una componente che corregge questi ultimi. A seconda della metodologia attraverso cui si implementano le regole CL, Rascu (2006) distingue due tipologie differenti di correttori:

- i correttori prescrittivi evidenziano tutte quelle strutture che sono consentite dai set di regole CL, ragion per cui la grammatica a cui queste sottendono delinea il modo in cui un testo deve essere scritto e prodotto;
- i correttori proscrittivi, invece, evidenziano tutte quelle strutture che sono proibite dai set di regole CL. La grammatica alla base di tali correttori definisce, pertanto, cosa, all'interno del testo, necessita di essere riscritto per conformarsi alle regole CL.

Di norma, nel settore, un correttore allo stato dell'arte è un ibrido che combina i due approcci, dove, da un lato, la componente proscrittiva evidenzia ciò che deve essere riformulato, mentre la componente prescrittiva indica come riformulare gli elementi problematici. A tal proposito, ad esempio, sono diverse le società di traduzione a progettare proprie utility di correzione automatica di regole applicata

all'output della machine translation.⁷ Tali utility, attualmente in fase di sviluppo, sarebbero in grado di applicare delle vere e proprie regole di post-editing, le quali andrebbero a implementare automaticamente le correzioni necessarie sulla base di set di regole prefissati, attraverso la computazione delle correzioni riscontrate nell'output della MT post-editato utilizzato per il training, applicando la cosiddetta distanza di edit.⁸ Appare chiaro come l'implementazione di funzionalità del genere costituirebbe un ulteriore strumento di perfezionamento dell'utilizzo della traduzione automatica dalle potenzialità notevoli.⁹

Per ragioni di brevità, in questa sede risulta impossibile elencare tutti i differenti set di regole CL esistenti, dal momento che ogni compagnia elabora il proprio per rispondere a esigenze e necessità di volta in volta differenti. Ciononostante, risulta opportuno elencarne alcuni tra i più comuni ed evidenziare quelli che risultano avere il maggiore impatto sulla qualità dell'output della machine translation. Dal momento che elaborare una classificazione delle differenti regole risulta altrettanto problematico e al di là degli scopi di questo studio, per comodità ci si servirà della classificazione elaborata da O'Brien (2003), la quale suddivide diverse regole, raccolte dai campioni di diverse compagnie che hanno aderito allo studio, in tre categorie: regole lessicali, regole sintattiche e regole testuali, a loro volta suddivise in strutturali e pragmatiche. Alla prima categoria fanno capo tutte quelle regole aventi lo scopo di influenzare la scelta del vocabolario da utilizzare e il significato delle parole stesse, al fine di rendere di più facile esecuzione l'individuazione dello scopo illocutivo della frase, un tipo di attività di sicura complessità per i sistemi MT che devono trovarsi a disambiguare, per esempio, tra l'espressione dell'obbligo o della possibilità (Rascu, *op. cit.*). A questa categoria appartengono inoltre regole che insistono sull'uniformazione dello spelling, l'uso di determinati quantificatori e congiunzioni, altresì proibendo, in linea di massima, un uso eccessivo di sinonimi. Come suggerisce il nome, tutte le regole di carattere sintattico rientrano nella seconda categoria sopraelencata, ovvero la concordanza tra soggetto e verbo, l'uso

⁷ Nel paragrafo seguente si descriveranno in maniera più dettagliata i meccanismi di funzionamento dei modelli di automatizzazione del post-editing, strettamente correlati alle regole CL.

⁸ Per approfondire il funzionamento della distanza di edit o distanza di Levenshtein, qui solo accennata brevissimamente, si veda www.levenshtein.net

⁹ Contenuto riadattato dal webinar tenuto da Tauyou il 6 gennaio 2016 come parte dei TAUS Translation Showcase Webinars, di cui si fornisce il link in bibliografia.

dell'ellissi e regole ben definite circa diatesi passiva, modo verbale e uso dei modali. Delle due sottocategorie testuali, infine, la prima mira a influenzare il layout stesso del testo, nonché la lunghezza delle frasi e dei paragrafi per non sovraccaricare il carico informativo degli stessi, mentre le regole di carattere pragmatico mirano a rendere il più esplicito possibile la funzione testuale attraverso l'organizzazione complessiva della frase e l'utilizzo dei verbi.

Delle regole sopraelencate, quelle più comuni risultano essere quelle che mirano a un uso standardizzato e il più chiaro possibile del vocabolario al fine di rendere il testo il più esplicito possibile. Inoltre, come risulta dal succitato studio (O'Brien, *op. cit.*), diverse regole sembrano fare divieto dell'uso di sintagmi nominali composti da più di tre nomi, altresì sconsigliando l'uso del gerundio.

Infine, da un'analisi degli standard di produttività, un uso accurato delle maiuscole e regole consolidate e standardizzate circa lo spelling sembrano avere l'impatto maggiore sulla qualità dell'output della machine translation. Infine, come è lecito aspettarsi, adottare un registro formale, evitando espressioni idiomatiche non troppo familiari ai sistemi MT, consente al sistema SMT di fare un uso coerente ed efficiente del materiale di training utilizzato, in cui figurano solo espressioni comuni e standard (Aikawa, *op. cit.*).

2.3.2 TECNICHE DI AUTOMATIZZAZIONE DEL POST-EDITING E *CONFIDENCE SCORE*

In precedenza, nel corso del capitolo in questione, si è fatto un breve accenno al notevole contributo che un traduttore può avere nell'evoluzione e potenziamento di un sistema MT. Effettivamente, la creazione di un sistema MT non consta solo della selezione dei giusti corpora e delle giuste regole grammaticali da implementare nel sistema stesso, ma bensì consta della partecipazione attiva del traduttore, ovvero di chi farà uso del prodotto finito. Di conseguenza, l'apporto che un traduttore può dare può spaziare dalla semplice selezione dei testi da tradurre a una partecipazione talmente attiva da fare del traduttore un vero e proprio sviluppatore del sistema MT stesso (Ryan, 1988).

Il contributo del traduttore/post-editor tuttavia non è necessariamente limitato a quanto detto finora. Diversi studi sostengono che, attraverso l'analisi del lavoro e

delle attività svolte da un post-editor, è possibile isolare ed evidenziare diversi trend ricorrenti nell'attività di correzione svolta dal post-editor, di cui servirsi per la creazione di moduli che di volta in volta automatizzino il processo di revisione andando ad affrontare automaticamente gli errori ricorrenti nell'output della machine translation, liberando il lavoratore di un carico di lavoro notevole e soprattutto ripetitivo non indifferente. Tali errori ricorrenti occorrono non solo a livello dello stesso documento ma tra diversi documenti nel corso del tempo, rendendone facile ed immediata l'individuazione da parte del post-editor navigato ed esperto. Analogamente al funzionamento delle TM per un traduttore, il quale grazie a queste non deve necessariamente tradurre sempre le stesse frasi, tali moduli consentirebbero di velocizzare il compito di un post-editor.

Suddetto procedimento può avvenire seguendo due differenti approcci. Il primo di questi, un approccio adattativo, prevede l'osservazione dell'attività di un post-editor e degli errori che vengono di volta in volta tradotti e la conseguente emulazione dell'attività dell'individuo. A tal proposito, nello studio di Allen e Hogan (2000) attraverso l'analisi di svariati set di file tradotti dall'inglese al francese e viceversa è stato infatti possibile individuare principi e trend ricorrenti adoperati nell'attività dei post-editor al fine di consentire a un modello di post-editing automatico (APE) di imparare le forme corrette a partire da suddetti set. Un dato interessante che tuttavia emerge da questo approccio è la possibilità che l'output della MT subisca processi di ipercorrettismo e che quindi tali modelli correggano espressioni del tutto accettabili nell'output. Ciononostante, lo studio ipotizza che, alla luce del guadagno produttivo ottenibile, eventuali correzioni non necessarie dell'APE avrebbero comunque effetti negativi marginali sulle prestazioni del post-editor (ivi).

Un secondo approccio, sebbene frutto di uno studio non più recente, prevede invece l'applicazione di un algoritmo in grado di facilitare l'attività del post-editor poiché mirato all'apprendimento delle forme corrette di quelle strutture che spesso generano problemi e necessitano dell'intervento di un post-editor. Per esempio, nello studio di Knight e Chander (1994) è stato possibile dimostrare come un modulo automatizzato di post-editing sia stato in grado di inserire automaticamente gli articoli "a", "an" e "the" in un testo inglese tradotto dal giapponese (ivi). Data la funzione logico-sintattica degli articoli all'interno di una frase, appare ovvio come le

informazioni apportate dal contesto, nonché le regole grammaticali applicabili, siano di vitale importanza per la corretta scelta dell'articolo appropriato. Codificare vaste informazioni grammaticali e pragmatiche tramite algoritmi è un'attività non del tutto facile, motivo per cui optare per un approccio che si basi sul supporto di dati risulta essere la scelta migliore, dato l'aiuto che esempi ed eccezioni di suddette regole possono fornire. Sulla base di un database di più di 400.000 sintagmi nominali estratti dal Wall Street Journal utilizzato come materiale di training, infatti, è stato possibile analizzare quali parole e combinazioni di parole, oltre alle relative caratteristiche, erano di maggiore aiuto nel predire l'articolo corretto. Ad esempio, nel database preso in esame, di fronte al nome composto "White house" in 238 occasioni l'articolo che figurava era "the", così come davanti modificatori quali "next" e "same". Tali dati hanno permesso di creare set di *features* dati dall'occorrenza, in più di un'occasione, di una data parola in una data posizione. Suddette feature erano caratterizzate dalla frequenza delle occorrenze, dalla distribuzione di "the" contro "a/an" e dalla distribuzione delle occorrenze in cui suddetta feature era assente. Per testare l'efficacia del post-editor automatico si è scelto di partire da quei sostantivi, aventi funzione di testa in un sintagma nominale, con il maggior numero di occorrenze nel materiale utilizzato nel training. Ad esempio, la parola "president" aveva un numero di occorrenze pari a 1420 e in queste, con una percentuale pari al 53.5%, l'articolo che la precedeva era il determinato "the". L'albero di decisione utilizzato per generare gli articoli per il sintagma nominale "president" ha generato un punteggio di accuratezza dell'89% (ivi). Ciò dimostra che con un numero sufficientemente grande di dati, è possibile generare alberi di decisione accurati tali per cui il sistema è in grado di scegliere, in maniera più o meno costantemente predicibile, l'articolo giusto, onde evitare quindi di incappare frequentemente negli stessi errori.

In una direzione differente, invece, si sono mossi invece ulteriori studi che mirano a munire il post-editor di una serie di strumenti che lo informino della qualità della MT e che al tempo stesso filtrino traduzioni ben fatte da traduzioni qualitativamente insufficienti. Tale approccio si basa sul modello della confidence estimation (Specia, 2009a, 2009b, 2010 e González-Rubio, 2015). Lo scopo della confidence estimation è quello di fornire dei punteggi riguardanti la stime della qualità dell'output della

machine translation, quando, a differenza di BLUE o NIST per esempio, non si dispone di una reference translation o, semplicemente, come nel caso di un post-editor monolingue, non si conosce la lingua source. Il vantaggio di un approccio di questo tipo è quello di poter essere direttamente implementato nel sistema SMT e di prevenire quindi la possibilità che il sistema suggerisca al post-editor traduzioni che necessitano di un intervento di correzione massiccio, poiché, in parole povere, tramite suddetti punteggi è il sistema stesso a indicare all'utente se quella traduzione è utilizzabile o meno (ivi). Lo scopo dei confidence score è quindi quello di indicare quanto il sistema sia sicuro della traduzione che sta suggerendo.¹⁰ I dati attraverso cui è possibile effettuare il training del sistema possono essere dati valutati automaticamente tramite BLEU o NIST o manualmente su una scala da 1 a 5, provenienti da diversi sistemi, domini e coppie di lingue, mentre la metodologia prevede l'applicazione, sui dati selezionati in base alla loro rilevanza, di un algoritmo in grado di restituire i suddetti confidence score. Attraverso questi il post-editor può infine servirsi di una stima della qualità grazie alla quale decidere se, in termini di guadagno produttivo, vale la pena post-editare la traduzione suggerita o tradurre il dato segmento da zero.

¹⁰ Contenuto estratto e rielaborato dalla *knowledgebase* consultabile sul sito internet di TAUS, di cui si fornirà il link in bibliografia.

2.4 IL POST-EDITOR: PROFILO E REQUISITI

In ultima battuta a quanto detto finora su standard di produttività del post-editing e misurazione dell'incremento di questi, pare opportuno riflettere anche sulla figura stessa del post-editor e sulla sua attività al fine di individuare quelle caratteristiche che contraddistinguono una buona performance da una carente e insufficiente. Per delineare queste caratteristiche è importante andare a evidenziare quelle abilità e peculiarità che fanno di un individuo un buon post-editor. Sebbene sia in parte lecito aspettarsi che le abilità di un buon traduttore siano equivalenti a quelle di un buon post-editor, sono diversi gli aspetti che smentiscono in maniera piuttosto immediata questa falsa credenza. Come riportato da Offersgaard (2008: 155), sono svariati i casi in cui, di fronte a un lavoro che richiedeva il post-editing di materiale pre-tradotto tramite machine translation, le performance di diversi traduttori si sono rivelate insufficienti a tal punto da non passare i test di Quality Assurance ed essere rigettate. Ciò è dovuto al fatto che le abilità richieste a un post-editor differiscono da quelle richieste a un traduttore, a eccezione di alcune comuni a entrambe le figure lavorative, ovvero quelle abilità richieste dal mercato e dalle agenzie di traduzione, quali ad esempio il rispetto delle guide di stile del cliente, nonché delle deadline. A queste vanno aggiunte la capacità di adottare la corretta terminologia, consolidata e coerente con guide di stile e materiale di riferimento, un'imprescindibile conoscenza del dominio oggetto della traduzione ed eccellenti abilità di revisione. Quali sono, allora, quei requisiti che differenziano il post-editor dal traduttore? Per quanto scontato possa sembrare, un bravo post-editor è colui il quale impiega meno tempo di un traduttore che traduce un dato segmento *from scratch*, utilizzando il più possibile l'output della MT, sapendo quando è possibile editare la traduzione suggerita dal sistema e quando è necessario ritradurre da zero.¹¹ Un buon post-editor è pertanto colui il quale è in grado di capire e decidere in un lasso di tempo di pochi secondi se quello che ha davanti è un prodotto buono sul quale è possibile lavorare per elevarne la qualità o se la traduzione presentatagli è da scartare interamente. Questo sembra essere uno dei compiti che richiede il maggior sforzo da parte dell'individuo, quello che più di tutti fa da discriminante tra il "buon" post-editor e il "cattivo" post-editor

¹¹ Parte dei succitati requisiti sono stati estratti e rielaborati dalle linee guida sul post-editing proposte da TAUS, di cui si fornirà il link in bibliografia.

(ivi: 156). Del resto, un post-editor efficiente è anche chi non apporta modifiche preferenziali, utilizzando per esempio sinonimi, fatti salvi i casi in cui questi sono stati esplicitamente indicati nelle istruzioni di progetto. A queste abilità essenziali O'Brien (2002) ne aggiunge diverse altre, integrate a quelle summenzionate. È interessante notare come, lì elencata tra le caratteristiche fondamentali quale *sine qua non*, figuri un atteggiamento positivo verso la machine translation, poiché parere comune è quello secondo cui chi viene forzato all'utilizzo del post-editing mostrerà un rendimento sempre e comunque inferiore a chi, per predisposizione propria, ha deciso di sposare volontariamente la causa del post-editing e della machine translation (ivi: 102). Tra le caratteristiche più prettamente teoretiche invece figurano una conoscenza vasta e approfondita della storia dello sviluppo della MT, nonché della teoria alla base del suo funzionamento. Infine, motivo per cui parte del capitolo in questione è stata dedicata alla loro breve introduzione e descrizione, completa il profilo una completa conoscenza dell'applicazione delle regole CL e dell'implementazione di moduli di automatizzazione del post-editing (ivi: 103).

Arrivati a questo punto, pare infine logico porsi l'interrogativo circa quale sia la figura professionale maggiormente portata per il post-editing e in che misura l'esperienza lavorativa personale incida sulle prestazioni del singolo. Sono differenti gli studi che mettono in correlazione anni di esperienza e produttività e qui di seguito verranno riportati i risultati di alcuni di essi.

In De Almeida e O'Brien (2010) le prestazioni di un campione composto da diversi post-editor, ciascuno avente alle spalle un numero diverso di anni d'esperienza, sono state analizzate e i risultati estrapolati hanno dimostrato che gli individui con il maggior numero di anni d'esperienza risultavano essere quelli più veloci, venendo a soddisfare pertanto un requisito fondamentale, dal momento che una maggiore velocità traduttiva è indispensabile a compensare tariffe per parola più basse. Coerentemente a questo risultato, gli individui con il minore numero di anni di esperienza erano quelli la cui velocità di produzione è risultata la minore.

D'altro canto, un dato in netta contraddizione con quanto detto finora sembra dimostrare che i post-editor con il minor numero di anni di esperienza sono quelli ad aver apportato il minor numero di cambiamenti e correzioni soggettive e non necessarie, mentre i post-editor più esperti hanno apportato un numero maggiore di

modifiche preferenziali, violando uno dei principi fondanti del post-editing. Ciò è stato probabilmente dovuto al fatto che una maggiore esperienza comporta un grado di indipendenza maggiore, mentre gli individui meno esperti devono essersi sentiti più vincolati alle linee guide del post-editing, con una conseguente maggiore “fiducia” della qualità dell’output della machine translation. Questo dato può anche fornire una spiegazione del perché diverse correzioni essenziali e necessarie siano state tralasciate da questi ultimi (ivi).

Da un punto di vista terminologico, invece, la tendenza mostrata dai post-editor con meno esperienza a sostituire diversi termini con sinonimi del tutto equivalenti, apportando quindi modifiche non necessarie, risulta essere coerente con quanto detto in precedenza, dimostrando quindi che maggiore esperienza equivale, con le dovute eccezioni, a maggiore accuratezza.

3. TRADUZIONE AUTOMATICA E POST-EDITING ALL'OPERA: ANALISI PRODUTTIVO-QUALITATIVA DELLE PRESTAZIONI DI MATECAT, SMARTCAT E SYSTRANET

3.1 TRADUZIONE AUTOMATICA E TOOL OPEN-SOURCE: ALCUNE CONSIDERAZIONI PRELIMINARI

All'inizio del primo capitolo si è brevemente accennato alla storia dell'evoluzione della machine translation e dei passi da giganti, non privi di qualche battuta di arresto, della tecnologia, i quali hanno portato al consolidamento della pratica del post-editing. Attualmente, nel mercato, sono molteplici i CAT tool che combinano utility più note e ormai assimilate, quali translation memory e termbase, a sistemi di traduzione automatica, per la maggior parte statistici. Questa spinta innovatrice verso una maggiore implementazione e assimilazione dei sistemi di traduzione automatica è un chiaro segnale di come la fiducia e il grado di apertura verso questa nuova tecnologia siano aumentati, nonché segnale che tanto l'utilizzabilità quanto la qualità della traduzione automatica sono arrivati a un livello tale da convincere software house e sviluppatori a investire su di essa.

Sebbene l'inizio degli anni Novanta segni l'inizio del periodo in cui la ricerca sulla SMT sia progredita di anno in anno e a livelli esponenziali, due avvenimenti hanno avuto il vanto di dare la spinta maggiore alla SMT, ovvero l'annuncio, nel 2011, che Google avrebbe reso disponibile, dietro sottoscrizione a un abbonamento, l'API di Google Translate, così da consentire l'implementazione di un sistema di traduzione automatica basato sul motore di ricerca numero uno al mondo nei programmi di traduzione assistita presenti sul mercato. Di lì a poco, Microsoft annunciò che, tramite modalità di abbonamento simili a quelle di Google, avrebbe reso disponibile l'API di Microsoft Translate. Eventi di questo genere sono un chiaro segnale della tendenza odierna a fare affidamento su tecnologie innovative come la SMT, con la conseguente proliferazione di programmi e strumenti appositi sul mercato. Della grandissima vastità di CAT tool fruibili dal vasto pubblico che se ne serve, una fetta piuttosto consistente è rappresentata da quei tool reperibili online gratuitamente, date le stime secondo cui circa tre quarti dell'utenza di Internet fa uso di programmi di traduzione online, gratuiti e open-source, complice la suddetta

maggior accessibilità a soluzioni di traduzione automatica. Un trend di questo tipo dovrebbe portare il lettore ad affermare, con un certo grado di certezza, che la proliferazione di tool gratuiti sulla rete è al tempo stesso causa e conseguenza della maggior utilizzabilità dei sistemi MT e di un approccio generalmente più aperto e positivo verso la machine translation. Tuttavia, è bene riflettere con obiettività e in termini di qualità/prezzo su pregi e difetti che derivano dal servirsi di prodotti che, di fatto, offrono soluzioni valide a costi nulli o ridottissimi, a confronto con programmi sviluppati da software house e sviluppatori leader del mercato. Del resto, non bisognerebbe mai dimenticare che oggi sono diversi i CAT tool che permettono di integrare e collegarsi a sistemi MT di parti terze, tramite plug-in o API key, offrendo un'ulteriore soluzione di lavoro che integra gratuitamente CAT tool a pagamento e MT di sviluppatori terzi.

Ovviamente, l'effettiva assenza di costi rende possibile l'utilizzo di tali prodotti in contesti che esulano da quello professionale e, pertanto, programmi di questo tipo risultano utili e soddisfacenti in contesti accademici e domestici, dove l'utente ha bisogno di prendere confidenza con le dinamiche di lavoro e l'impostazione dei progetti di traduzione condotti tramite l'utilizzo di CAT tool, a fronte del fatto che le licenze messe a disposizione da programmi a pagamento sono limitate nel tempo e nelle funzionalità, nonché non del tutto gratuite in alcuni casi. Data la gratuità dei programmi open-source, d'altro canto, questi sono per forza di cose costretti, salvo alcune eccezioni, ad ambienti amatoriali e domestici, al contrario dei software a pagamento, largamente più usati in contesti professionali e lavorativi. Ciononostante, molti di questi programmi open-source, in virtù delle loro comunque valide potenzialità, si rivelano essere dei validi punti di partenza o stadi intermedi in un più grande progetto di traduzione, data la possibilità per diversi di questi di elaborare e utilizzare formati di scambio di translation memory (TMX). Inoltre, dal momento che uno dei pochi requisiti, se non l'unico, è disporre di un browser web, un tool open-source è utilizzabile in virtualmente tutti i sistemi operativi, siano essi Windows o Mac OS.

D'altro canto, per forza di cose, programmi e applicazioni disponibili gratuitamente online mostrano le loro carenze quando costretti a lavorare su progetti più complessi che richiedono la partecipazione di più traduttori, ognuno con proprie

lingue di lavoro. Inoltre, caratteristica comune a diversi tool open-source è la mancanza o scarso rendimento, laddove presenti, di applicazioni per la gestione del materiale traduttivo di supporto, quali strumenti di gestione terminologica e di memorie di traduzione, dizionari specifici di dominio, nonché strumenti per revisione e QA, ragion per cui alcuni di questi tool si servono di applicazioni terze progettate indipendentemente dal CAT tool open-source o miratamente per ovviare a determinate pecche o mancanze di funzionalità.

A discapito dei sopra descritti punti deboli, alcuni dei quali più problematici di altri, in virtù dei comunque apprezzabili vantaggi offerti dai CAT tool open-source, una grossa fetta del mercato traduttivo sembra essersi concentrata, con attenzione e impegno sempre maggiori, sulla ricerca e progettazione di suddetti tool pensati per un bacino di utenza considerevolmente più ampio di quello delle software house proprietarie, al punto che alcuni di questi hanno dimostrato di possedere le potenzialità per concorrere a tutti gli effetti nel mercato dei tool per traduzione assistita, automatica e post-editing. Bisognerebbe chiedersi se questa tendenza può essere interpretata come il riflesso di un crescente interesse nella machine translation e nel post-editing, i quali, da facili bersagli di scetticismo e diffidenza, sembrano invece godere ora in maniera sempre più importante di posizioni di rilievo nell'industria, a tal punto da divenire fonti di guadagno e investimento. Non a caso, diversi tool gratuiti di questo tipo sono il prodotto di compagnie nate come agenzie di traduzione e solo in seguito espansive come sviluppatrici di CAT tool propri, resi disponibili gratuitamente o dietro pagamento mensile o annuale di modiche cifre, con l'evidente scopo di fidelizzare il maggior numero di clienti, siano essi professionisti del settore o meno.

A tal proposito, scopo di questo capitolo sarà analizzare in maniera obiettiva e con dati certi alla mano le prestazioni di diversi tool gratuiti o semi-gratuiti pensati per questa porzione di utenti, prevalentemente alle prime armi e interessata in questa tipologia di CAT tool, al fine di ipotizzare un possibile utilizzo di queste applicazioni e magari predirne una loro futura implementazione in contesti lavorativi o professionali; in caso contrario, individuarne le debolezze intrinseche e ammetterne i limiti, isolando l'utilizzo di suddetta tecnologia a contesti meno complessi di quelli imposti dal mercato.

3.1.1 CARATTERISTICHE E FUNZIONALITÀ DI MATECAT

Nel giugno 2015, ad Angers, Francia, si sono tenuti diversi cicli di conferenze su CAT tool, traduzione assistita e automatica.¹ Tra i diversi sviluppatori che si sono presentati per condividere idee e proporre proprie versioni dei loro tool, figuravano diversi enti universitari di ricerca e agenzie di traduzione, tra cui la ben nota Translated. Quest'ultima ha presentato il CAT tool open-source di propria progettazione, Matecat.² La nascita di Matecat, acronimo di Machine Translation Enhanced Computer Assisted Translation, ha origini piuttosto recenti: nato come progetto sviluppatosi nell'arco di tre anni grazie ai fondi stanziati dal Settimo Programma Quadro (7PQ) dell'Unione europea, attraverso varie versioni beta promosse grazie a intensive campagne promozionali e sfruttando eventi di grande attenzione pubblica, quali ad esempio il Campionato mondiale di calcio 2014, occasione in cui venne offerta la possibilità di testare le potenzialità del tool nella traduzione di un breve testo avente per tema proprio la competizione calcistica, dal 2014 Matecat si è imposto sul mercato quale tool gratuito e open-source per sfruttare al meglio le potenzialità di traduzione automatica e post-editing. L'obiettivo di Matecat e del team dietro la sua progettazione è quello di fornire uno strumento gratuito e open-source in linea con bisogni e necessità degli utenti, offrendo un alto livello di integrazione della machine translation e incoraggiando gli utenti alla creazione di un ambiente collaborativo, agevolato da un alto livello di accessibilità online. Il motore di traduzione automatica utilizzato è il ben noto MyMemory. Implementabile nella catena di lavoro e nel processo traduttivo, il sistema MT di Matecat consente di incrementare la propria produttività grazie all'apprendimento e aggiornamento in tempo reale delle correzioni che l'utente implementa durante lo svolgimento del proprio lavoro.³

L'interfaccia utente risulta essere estremamente semplice e pulita, e già da questa è possibile vedere l'attenzione posta nel creare un strumento in grado di attrarre il maggior numero di utenti, più o meno esperti, facendo leva sui vantaggi apportati

¹ Per una panoramica completa dell'incontro, si consulti il sito internet dell'evento, <http://www.tao2015.org>

² www.matecat.com

³ Contenuto riadattato dal webinar "Gengo & Translated", disponibile sul sito Internet di TAUS e tenutosi il 2 ottobre 2013, durante la prima dimostrazione ufficiale di Matecat. Per la visione integrale del webinar, si veda il link disponibile in bibliografia.

dall'utilizzo della machine translation. Nella prima schermata del tool, raggiungibile semplicemente digitando l'indirizzo internet, dopo aver impostato le lingue di lavoro, aver caricato i file da tradurre e impostato il dominio dei documenti, grazie al quale il tool può restringere la ricerca all'interno dei dati usati per alimentare il sistema MT (Schwenk *et al.*, 2015), Matecat provvederà ad analizzare i documenti da tradurre, restituendo una schermata come quella mostrata in figura 3.1.

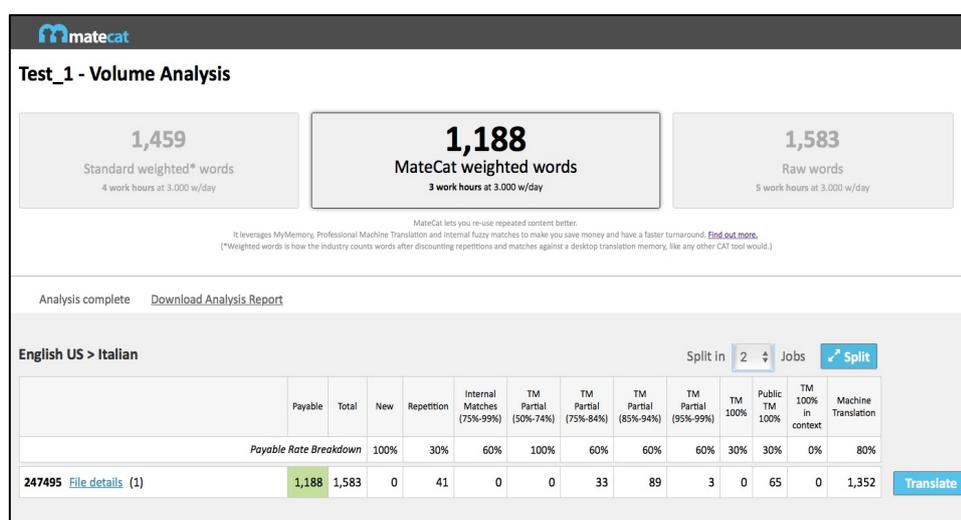


Figura 3.1 Schermata dell'analisi dei documenti da tradurre in Matecat. Matecat restituisce il numero totale di parole da tradurre senza il calcolo di ripetizioni, fuzzy match o context match e il relativo quantitativo di tempo necessario, a un ritmo di 3000 parole al giorno (riquadro a destra), il numero pesato di parole da tradurre a fronte dei diversi tassi per parola applicati a ripetizioni e fuzzy match (riquadro a sinistra) e infine il numero di parole da tradurre pesate secondo le modalità utilizzate da Matecat.

Come si evince dalla figura, la funzione principale dell'analisi è quella di mostrare il considerevole guadagno, in termini di tempo necessario, grazie all'utilizzo della MT, in virtù di un conteggio delle parole da tradurre che non è altro che la combinazione di memorie di traduzione e machine translation, dove la percentuale applicata a quest'ultima varia tra l'80 e il 90 per cento, calcolata per ogni coppia di lingua in base all'utilità che la traduzione automatica ha avuto sull'ultimo milione di parole tradotte nella coppia di lingue presa in considerazione, partendo dal presupposto che minori sono le modifiche apportate dal traduttore ai suggerimenti della machine translation, maggiore è la sua utilità.

Un'altra caratteristica su cui è posta moltissima attenzione e cura è la possibilità di dividere i file da tradurre tra più traduttori in maniera estremamente facile, a

favore di un'ormai sempre più frequente tendenza all'outsourcing, fortissimamente promosso da Matecat. Accorgimenti di questo tipo rendono assolutamente conseguenziale individuare una cura minuziosa nella creazione di un tool marcatamente portato per il project management, tramite il quale predisporre il progetto, gestirlo in remoto e inviarlo ai diversi traduttori, i quali avranno solo bisogno dell'URL di progetto, creato al momento della creazione dello stesso. Questa forte componente organizzativa, obiettivamente meno pensata per potenziare le attività di traduzione e post-editing ma più quella di project management, è evidente anche in un'altra serie di funzionalità, tra cui spicca quella che permette di creare un vero e proprio ordine di traduzione, nel quale figurano tempi, costi di traduzione e altre informazioni rilevanti per l'eventuale cliente, come mostrato in figura.

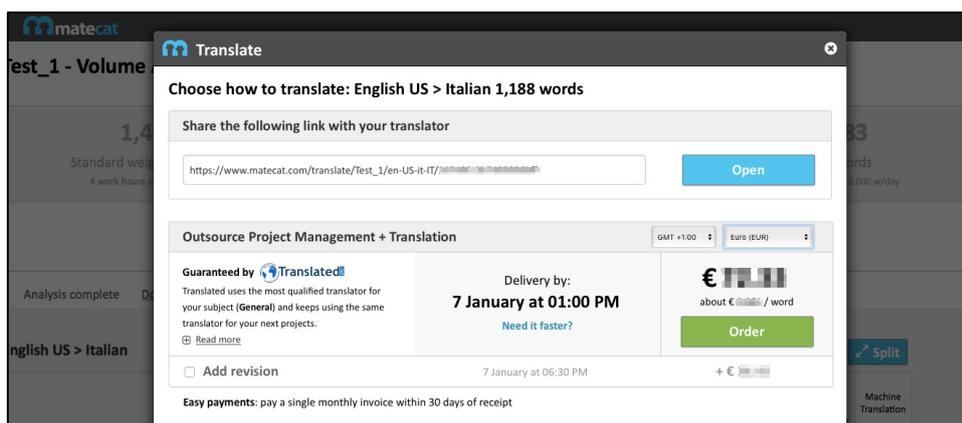


Figura 3.2 Schermata della richiesta di ordine di traduzione in Matecat. Il link in cima consente di condividere il progetto tra più traduttori.

Al di là di queste funzionalità, quelle oggetto di maggiore interesse in questa sede sono quelle che hanno a che fare con le attività di traduzione e post-editing vere e proprie. L'interfaccia dell'editor di Matecat si dimostra intuitiva ed estremamente semplice da utilizzare, come appare in figura 3.3, dove i suggerimenti di traduzione automatica e TM pubbliche vengono visualizzati sotto ogni coppia di segmenti.

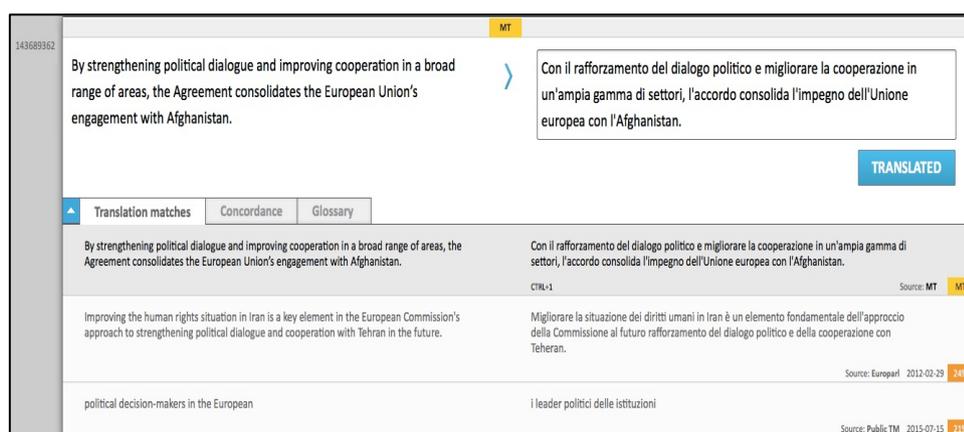


Figura 3.3 Schermata editor di Matecat. Al di sotto di ogni coppia di segmenti vengono visualizzati i suggerimenti della machine translation e di eventuali TM pubbliche disponibili, con le relative percentuali di matching.

Analogamente alla schermata dell'editor, la schermata di revisione (Figura 3.4) si presenta in maniera del tutto simile, con l'unica eccezione della sezione dedicata ai parametri di revisione, attraverso i quali è poi possibile creare un report della qualità globale della traduzione.

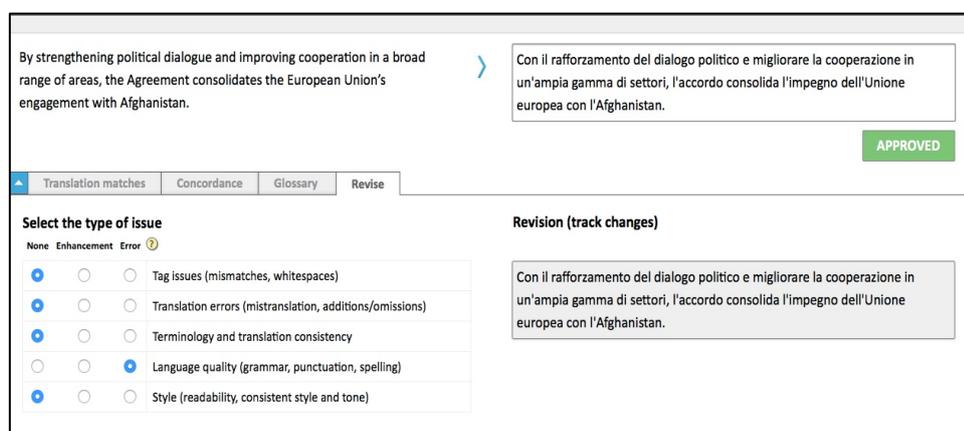


Figura 3.4 Schermata di revisione di Matecat. Tramite i dati forniti dal revisore, Matecat calcola il punteggio medio attribuibile alla qualità della traduzione. Ad ogni categoria viene attribuito un punteggio in base al numero di errori riscontrati e il punteggio finale è la media dei punteggi di ogni singola categoria.

Un dato interessante al riguardo emerge dal fatto che eventuali carenze nella personalizzazione dei parametri di revisione, apparentemente standard e non molto dettagliati in termini di diverse tipologie di errore, sono compensate dalla possibilità di implementare le funzionalità di tool sviluppati da parti terze, come ad esempio TAUS DQF (Dynamic Quality Framework).

3.1.2 CARATTERISTICHE E FUNZIONALITÀ DI SMARTCAT

Impostosi nel mercato nei primi mesi del 2015, SmartCAT⁴ è un CAT tool progettato dall'agenzia di traduzione, con sede a Mosca, Abbyy LS. Disponibile sottoforma di servizio cloud-based dotato di alti livelli di flessibilità e personalizzazione, SmartCAT offre una vasta gamma di servizi per la traduzione assistita e automatica, con una cura particolare per l'esperienza lavorativa e professionale per traduttori e localizzatori freelance. Al contrario di altri tool interamente gratuiti, SmartCAT consente l'utilizzo gratuito di gran parte delle sue funzionalità a traduttori freelance, mentre per compagnie e agenzie di traduzione l'accesso alle sue funzionalità è garantito previo pagamento di un abbonamento. Diversamente, l'implementazione della machine translation nel tool è soggetta al pagamento di un quota che varia a seconda del numero di pagine tradotte attraverso l'utilizzo della MT, dal momento che SmartCAT si serve di sistemi di traduzione automatica di parti terze, ovvero Google e Microsoft Bing. Qualora il progetto di traduzione sia impostato sulla traduzione da e verso lingue slave, il sistema MT associato è quello di default ABBYY MT.

Una funzionalità del tutto innovativa, anch'essa disponibile a pagamento, è l'OCR (Optical Character Recognition), ovvero la possibilità di tradurre il testo contenuto in immagini acquisite o in file PDF, JPG e altri formati. Tale funzionalità è facilmente raggiungibile ed eseguibile nelle impostazioni di progetto durante la creazione di quest'ultimo, come mostrato in figura 3.5.

⁴ www.smartcat.ai

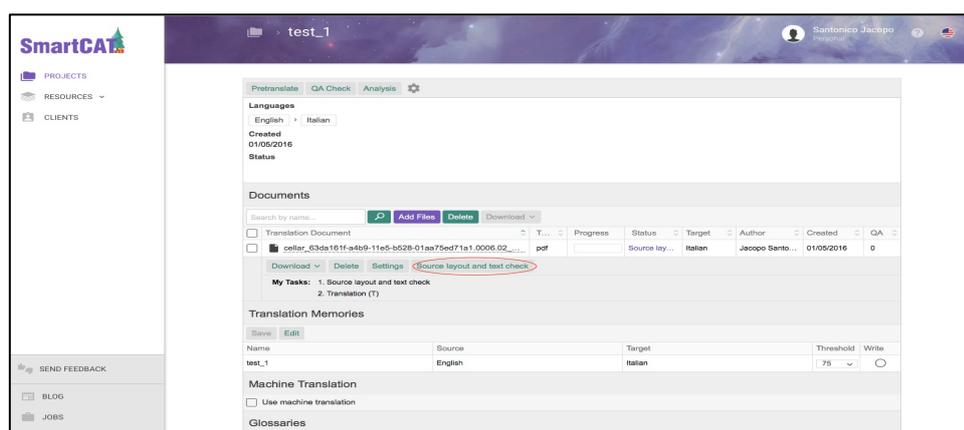


Figura 3.5 Schermata delle impostazioni di progetto di SmartCAT. Prima di tradurre i documenti del progetto, per ognuno di questi sono disponibili diverse funzionalità tra cui l’analisi del contenuto grafico e del layout del testo source, al fine di estrarre eventuali elementi che devono essere tradotti, tramite il comando Source layout and text check.

A differenza delle memorie di traduzione e dei glossari, interamente personali e personalizzabili, così come per la machine translation anche l’utilizzo di dizionari di dominio specifico è possibile grazie ad applicazioni terze, rese pubbliche per l’utilizzo e pertanto non modificabili secondo le proprie esigenze. SmartCAT, infatti, si serve dei dizionari di Abbyy Lingvo, disponibili gratuitamente ma limitati a poche lingue, principalmente da e verso lingue slave.

Al di là delle funzionalità standard che ogni CAT tool che si rispetti dovrebbe avere, quali per esempio termbase, memorie di traduzione, *autopropagation* e *concordance search*, presenti pertanto anche in SmartCAT, è evidente come una cura maggiore sia stata posta nello sviluppo del comparto di revisione e QA, grazie a una vasta gamma di opzioni personalizzabili fin nei minimi dettagli, grazie alle quali è possibile assicurare un controllo della qualità del testo tradotto approfondito e dettagliato. Tra le opzioni personalizzabili, mostrate in figura 3.6, spicca sicuramente la possibilità di evidenziare eventuali incoerenze nella traduzione di termini specifici, nonché l’inosservanza di terminologia consolidata e specifica del cliente di progetto.⁵

⁵ Contenuto estratto dall’esperienza diretta del tool e in parte rielaborato dal webinar “Translation Technology Showcase Webinar” del 6 maggio 2015, di cui si fornirà il link in bibliografia.

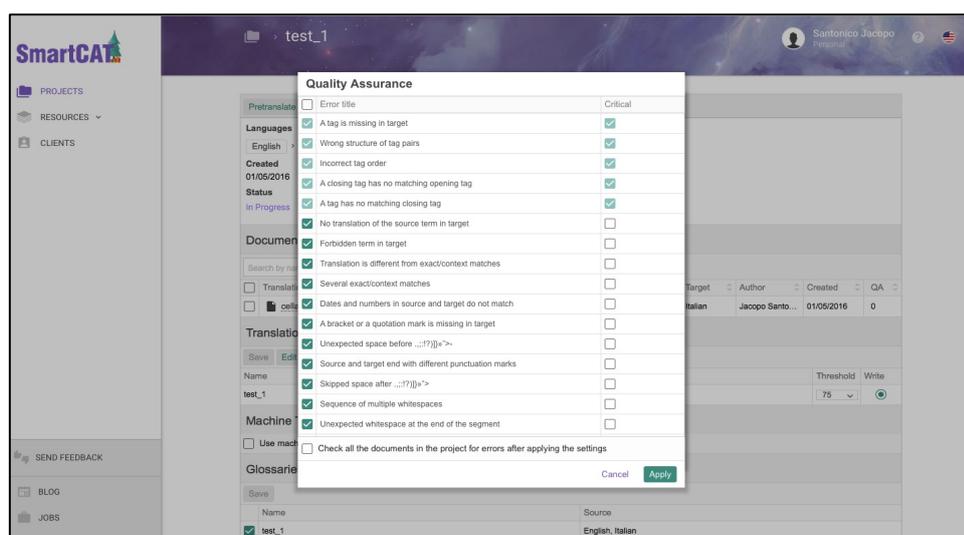


Figura 3.6 Lista delle opzioni attivabili nel menù Quality Assurance di SmartCAT. Per ogni errore è possibile impostare il grado di importanza di questi: gli errori designati come critici determineranno l'impossibilità di confermare il segmento in cui è presente suddetto errore.

3.1.3 CARATTERISTICHE E FUNZIONALITÀ DI SYSTRANET

Systran, ancora oggi, risulta essere uno dei primi tool votati alla machine translation. Come precisato nel primo capitolo, il suo primo utilizzo risale al 1970, mentre la sua prima applicazione quale tool di traduzione automatica dei documenti della Commissione europea risale al 1976. Oggi, Systran offre una vasta gamma di prodotti per la MT, siano essi prodotti desktop, prodotti server o disponibili online. Tra quest'ultimi rientrano diversi tool online gratuiti per la traduzione in tempo reale di siti Web e file. Si tratta di programmi le cui funzionalità sono limitate alla semplice traduzione e quindi privi di funzionalità aggiuntive quali memorie di traduzione, termbases o QA checker. Ciononostante, questi tool offrono prestazioni degne di un qualunque altro CAT tool, forti del sistema di traduzione automatica di Systran, un sistema che combina rule based e SMT. Di questi, in questa sede l'interesse maggiore sarà rivolto verso SYSTRANet,⁶ un CAT tool per la traduzione automatica interamente gratuito. Come mostrato in figura 3.7, l'interfaccia utente risulta piuttosto scarna, dotata delle funzionalità essenziali per la traduzione di testi, siti internet e file Word, RTF, HTML, HTM e TXT.

⁶ www.systranet.com

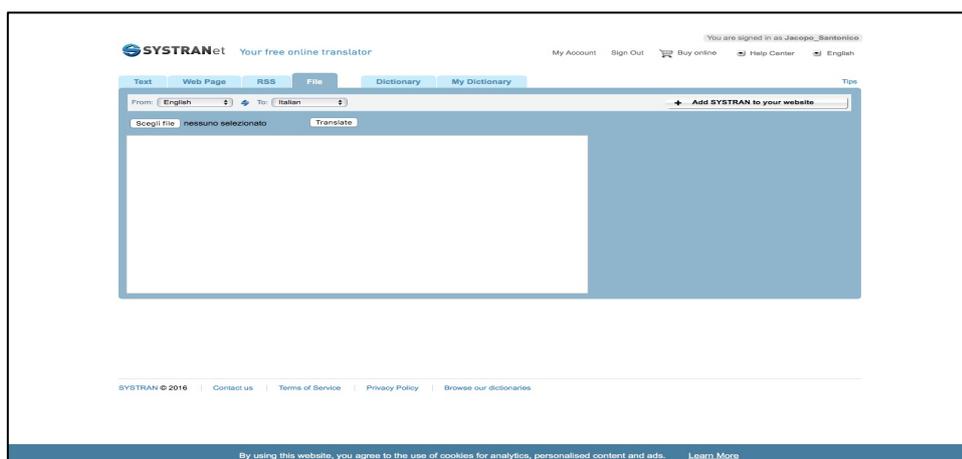


Figura 3.7 Interfaccia utente di SYSTRANet. I pannelli in alto a sinistra consentono la traduzione di testi di breve lunghezza, pagine Web, feed RSS e file.

Il numero di formati su cui può lavorare SYSTRANet risulta piuttosto limitato, soprattutto a confronto con gli altri due tool. L'impossibilità di lavorare su file PDF, per esempio, costituisce un limite non indifferente, se si considera l'ingente numero di testi prodotti e poi distribuiti oggi in questo formato. Pertanto, qualora si scelga di utilizzare un tool come SYSTRANet, è importante ragionare anche e soprattutto in termini di utilizzabilità del tool rispetto ai formati previsti dal progetto, poiché eventuali incompatibilità potrebbero risultare in tempi ulteriori di preparazione e conversione dei documenti da tradurre. Come si avrà modo di vedere, infatti, quello del formato sarà una questione che condiziona in maniera importante il lavoro di prova che si condurrà sui tool presi in esame.

Sebbene SYSTRANet non disponga di un QA checker e di funzionalità che consentano la revisione dei testi tradotti, il tool compensa in parte questa mancanza con la possibilità di interrogare i dizionari online SYSTRAN e Larousse (figura 3.8), nonché la possibilità di crearne di propri, aggiungendo manualmente i termini desiderati, ai quali è possibile aggiungere informazioni utili, quali la categoria grammaticale di appartenenza o, alternativamente, lasciare che sia il sistema a riconoscerla.

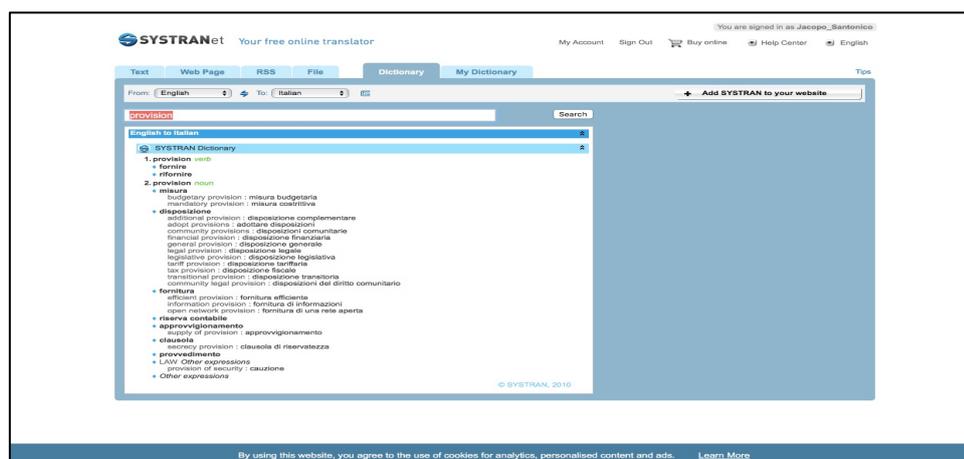


Figura 3.8 Interfaccia Dictionary di SYSTRANet. Selezionando lingua source e lingua target, SYSTRANet interroga i dizionari SYSTRAN e Larousse e fornisce eventuali traduzioni, differenziate per categoria grammaticale e dominio d'applicazione del lemma.

3.2. LA TIPOLOGIA TESTUALE ESAMINATA

In modalità più o meno equivalenti tra loro, sono diverse le tipologie di testi e domini che dimostrano di restituire prestazioni sufficienti e valide quando sottoposti a sistemi di traduzione automatica. Tra questi, un dominio in particolare è spesso oggetto di studi e prove dimostrative per testare potenzialità e qualità di un tool di traduzione automatica, ovvero il dominio dei testi legali. Coerentemente a suddetti studi, pertanto, al fine di restituire risultati correlabili con quelli già mostrati nella letteratura in materia, la tipologia testuale oggetto di questa analisi contrastiva sarà proprio quella giuridico-legale e, più precisamente, nell'ambito comunitario.

Da un punto di vista prettamente linguistico pragmatico, suddetta tipologia testuale si distingue dalle altre in virtù di caratteristiche e proprietà ricorrenti non individuabili in altre, o quantomeno individuabili in misura minore. La letteratura in materia è tra le più vaste, spaziando dalla linguistica testuale e pragmatica alla sociolinguistica. Pertanto, senza entrare troppo nel dettaglio e confondere il lettore con principi che esulano dalla materia oggetto del presente studio, risulta comunque necessario indagare su alcune delle summenzionate caratteristiche dei testi legali, al fine di comprendere meglio i fenomeni scaturenti dalla traduzione automatica di questi ed elaborare, di conseguenza, ipotesi e possibili strategie risolutive.

In virtù di un numero consistente di tratti distintivi, sono diverse le tassonomie che hanno per obiettivo quello di classificare i differenti testi in tipologie testuali ben

differenziate tra loro. Al fine di distinguere i differenti tipi di testo è opportuno individuare le diverse funzioni illocutive del testo, ovvero lo scopo per cui è stato scritto il testo, e il vincolo interpretativo che intercorre tra emittente e destinatario (Sabatini, 1999: 142).⁷ Per vincolo interpretativo si intende il maggiore o minore grado di esplicitezza tramite cui vengono trasmesse le informazioni del messaggio, per regolare il livello d'univocità dell'interpretazione di queste. A seconda della tipologia testuale e dell'intento comunicativo, l'autore del testo vorrà regolare in maniera più o meno esplicita la libertà interpretativa da parte del destinatario. Nella fattispecie, i testi legali, ovvero testi costrittivi normativi che mirano a fornire norme di comportamento inequivocabili, sono testi altamente vincolanti, i quali richiedono una conoscenza ben determinata del contenuto a cui fanno riferimento, motivo per cui questi fanno frequentissimamente riferimento a un sistema di testi connessi e richiamati; tramite suddetti richiami intertestuali accrescono la loro rigidità e il grado di esplicitezza (ivi: 146-148). Tali esigenze si riflettono anche sul piano sintattico e sulle scelte stilistiche effettuate. Infatti, per garantire il massimo grado di esplicitezza, un testo legale prediligerà l'esplicitazione del soggetto, costantemente presente tramite ripetizioni nominali o sostituenti pronominali, e di tutti gli argomenti del verbo, a discapito di altri artifici retorici quali l'ellissi. La struttura delle frasi di un testo legale è pertanto estremamente standardizzata e rigida, in modo da ridurre al minimo possibile l'ambiguità semantica dell'informazione, negando qualunque inferenza da parte del lettore, atteggiamento che si riflette ulteriormente nella quasi totale assenza o quantomeno rarissima presenza di connettivi frasali (ivi: 152-161). Da un punto di vista traduttivo, la considerevole rigidità del testo e la quasi totale assenza di costruzioni che possono risultare ambigue facilità il lavoro di un sistema di traduzione automatica, dal momento che questo difficilmente si imbatte in frasi fortemente connotate dal punto di vista semantico. D'altro canto, i testi legali risultano essere fortemente connotati da un punto di vista più squisitamente stilistico e, complici i frequenti rimandi intertestuali, fanno riferimento a sovrastrutture comuni a tutti i testi legali, motivo per cui garantire coerenza stilistica, nonché terminologica, può risultare un compito talvolta arduo.

⁷ Per una maggiore comprensione del concetto di vincolo interpretativo e, più in generale, del concetto di patto comunicativo alla base della tassonomia di Sabatini, si consiglia la lettura integrale del suo scritto, qui brevemente citato, del quale si fornirà il rimando bibliografico nella sezione Bibliografia.

3.2.1 L'ITALIANO COMUNITARIO: L'EUROLETTO

La varietà di testi legali di ambito comunitario oggetto di traduzione nelle istituzioni dell'Unione risulta essere tra le più vaste e categorizzabile in base alla natura più o meno giuridica degli stessi e in base al peso dettato da una presenza maggiore o minore della terminologia comunitaria in essi contenuti. In quest'ottica pertanto è possibile individuare atti legislativi e non legislativi, i quali rientrano nella serie L della Gazzetta ufficiale dell'Unione europea, nonché risoluzioni, raccomandazioni, pareri, comunicazioni, atti preparatori, informazioni e avvisi, i quali rientrano nella serie C della Gazzetta ufficiale. Tra gli atti legislativi figurano regolamenti, direttive, decisioni e bilanci, mentre accordi internazionali, alcuni regolamenti, direttive e decisioni, raccomandazioni e orientamenti rientrano negli atti non legislativi. I documenti che rientrano nel più grande contenitore della serie C, invece, sono tra i più differenziati tra loro e tra di essi si possono trovare risoluzioni, raccomandazioni, pareri, dichiarazioni comuni, comunicazioni e informazioni provenienti da istituzioni, organi e organismi dell'Unione europea, nonché procedimenti amministrativi e giurisdizionali (Unione europea, 2011: 21-37).⁸

Questa considerevole varietà impone strategie traduttive di volta in volta differenti ma comunque votate a un atteggiamento prevalentemente conservativo nei confronti di lessico, stile e terminologia, dal momento che ogni scelta traduttiva porta con sé rilevanza considerevole sul piano politico. Per tale motivo, quindi, con l'obiettivo di limitare al massimo gli interventi traduttivi creativi, spesso e volentieri quella della traduzione in ambito comunitario si rivela essere una pratica fondamentalmente ridotta alle giuste ricerche in banche documentali e terminologiche (Cosmai, 2007: 97). La forte rigidità stilistica a cui si è fatto riferimento precedentemente condiziona quindi l'atto traduttivo, il quale è vincolato a formule standardizzate e terminologia tecnica. Inoltre, gli atti comunitari sono soggetti a un modello strutturale ben determinato, costituito da titolo, preambolo, dispositivo o articolato e formula finale (ivi: 98). Questa struttura impone la massima rigidità, soprattutto nella parte iniziale, ovvero nel preambolo, sebbene, in determinati documenti come regolamenti e direttive, tale rigidità si estenda anche al dispositivo, in maniera tale che, per quanto

⁸ Le suddette informazioni, nonché i criteri redazionali presi a modello in questo studio, sono reperibili nel Manuale interistituzionale di convenzioni redazionali, di cui si fornirà il rimando bibliografico nella sezione Bibliografia.

la lingua di partenza presenti lievi differenze rispetto alla formulazione consolidata, il traduttore dovrà comunque attenersi al modello consolidato della lingua di arrivo (ivi: 100). Inoltre, data l'alta intertestualità di tali documenti, la terminologia consolidata ha valore cogente nel momento in cui fa riferimento a testi connessi e a essi richiamati.

La struttura fortemente standardizzata dei testi comunitari fa sì che i traduttori delle istituzioni europee si affidino a strategie traduttive ricorrenti e consolidate a tal punto da aver dato vita nel corso degli anni a un codice linguistico proprio dei testi prodotti in seno all'UE, codice che con il tempo ha assunto l'identità e nome di euroletto. Tutte le lingue dei Paesi membri dell'Unione godono dello statuto giuridico di lingue ufficiali, motivo per cui nel regime linguistico dell'Unione non si parla mai di testo originale, testo tradotto e traduzione, ma bensì di testi paralleli e corelazione, ovvero una stesura comune e parallela delle diverse versioni linguistiche. Per tale motivo, l'euroletto si configura come un sottocodice considerevolmente radicato nel contesto comunitario, nato come naturale conseguenza di una continua operazione traspositiva interlinguistica. Infatti, nelle versioni tradotte sono rintracciabili le modalità espressive del testo originario, tanto è forte il legame tra i due testi, motivo per cui si può parlare di un testo ibrido dotato di un modello linguistico presente in maniera trasversale nei diversi testi di arrivo (Mori, 2008: 533-535).

Da un punto di vista linguistico, la continua influenza delle lingue europee, le une sulle altre, ha portato alla genesi di un codice, fatto di forestierismi, neologismi ed europeismi, atto a esprimere i nuovi concetti introdotti in seguito alla nascita di un'identità sovranazionale imponente quanto quella comunitaria. Un aspetto che ha assunto grande rilevanza nell'italiano comunitario è l'addensarsi di europeismi, siano essi sintagmatici come comunità europea o Stati membri, europeismi di forma quali sono tutte le sigle e gli acronimi di cui abbondano i documenti comunitari, o gli europeismi di senso, ovvero termini preesistenti che, in seguito a un processo di risemantizzazione, hanno subito un arricchimento a livello semantico (ivi: 538-539).

Da un punto di vista sintattico, invece, tramite l'osservazione di testi comunitari tradotti dall'inglese, lingua veicolare per eccellenza dell'UE dopo l'adesione della Gran Bretagna nel 1973, è possibile isolare ed evidenziare una serie di strategie

traduttive ricorrenti che insieme danno vita a veri e propri pattern stilistici tipici dell'italiano comunitario. Da una prima analisi, infatti, emerge una forte tendenza alla sovraesplicitazione di legami logico-sintattici lasciati impliciti nella lingua inglese, oltre a una maggiore tendenza all'utilizzo di sintagmi preposizionali in luogo del soggetto inglese (Mori, 2002: 57). Tale tendenza si riscontra anche nell'inversione frequente tra soggetto e verbo, possibile nella lingua italiana e non ammessa in quella inglese, motivo per cui si assiste spesso alla tematizzazione del verbo dislocato in posizione iniziale, a favore quindi di una prospettiva funzionale della frase (ivi: 60). Inoltre, nei testi comunitari italiani è possibile rintracciare una generale preferenza per la diatesi passiva e il si impersonale (ivi: 61), nonché un consistente addensamento di nomi, dovuti alla traduzione delle frequenti -ing form presenti nella lingua inglese (ivi: 58). Infine, nei testi comunitari italiani è possibile osservare la realizzazione della modalità deontica della frase attraverso la resa in italiano dei frequenti "will" e "shall" con il presente prescrittivo (*ibid.*).

Sebbene i fenomeni analizzabili siano molteplici e la premessa appena fatta ne ha solo brevemente indagati alcuni, questa sarà di fondamentale importanza per ritrovare suddetti fenomeni nel testo comunitario del quale si andrà a valutare la traduzione prodotta dai sistemi MT, al fine di valutare le prestazioni di quest'ultimi rispetto a tali fenomeni.

3.3 COME VALUTARE LA QUALITÀ DELLA MACHINE TRANSLATION: UNA METODOLOGIA DI LAVORO

Al fine di condurre un'analisi obiettiva e dettagliata delle prestazioni dei summenzionati CAT tool sarà opportuno approntare una metodologia di lavoro e studio che affronti tale compito dai diversi punti di vista cruciali all'analisi e che restituisca dati certi tramite i quali poter avanzare ipotesi e postularne l'effettiva veridicità. A tal proposito, sarà pertanto necessario testare le potenzialità dei tre tool nella traduzione automatica di uno stesso testo al fine di individuare pattern ricorrenti comuni o divergenti tra i tre tool.

Il testo comunitario scelto per testare le potenzialità dei tre tool è una proposta congiunta di decisione del Consiglio,⁹ facilmente reperibile sul portale del diritto comunitario Eur-lex. Il motivo per cui è stato scelto tale documento è semplice: al di là della tipologia di atto, non del tutto rilevante ai fini della presente analisi, suddetto documento è stato scelto poiché manchevole del testo parallelo italiano, la cui eventuale presenza in rete avrebbe potuto falsare il lavoro dei tool e di conseguenza dei dati. Ciononostante, documenti aventi struttura e argomento simile, ovvero un accordo di cooperazione sul partenariato e lo sviluppo, sono reperibili e consultabili anche nella loro versione italiana, alla luce della succitata standardizzazione, ricorrenza e cristallizzazione di strutture sintattiche e stile dei testi comunitari. Questo dato sicuramente influenzerà le prestazioni dei tool, in modalità che si andranno ad analizzare a seguire.

Al fine di condurre un'analisi obiettiva delle prestazioni di Matecat, SmartCAT e SYSTRANet sarà opportuno approntare una metodologia di valutazione che tenga conto dei differenti aspetti che insieme concorrono a definire in maniera più o meno positiva o negativa la qualità dell'output dei tool, nonché le potenzialità e i limiti intrinseci dei suddetti. Come ampiamente trattato nel capitolo precedente, l'analisi della qualità di un testo tradotto tramite MT deve rispondere a parametri ben definibili e in grado di restituire risultati coerenti quando applicabili a differenti sistemi. Per tale motivo, prima servirsi di tool in grado di compiere questa valutazioni in maniera precisa e dettagliata, ovvero i QA checker, sarà utile condurre un'analisi preliminare del testo tradotto automaticamente applicando i parametri di macrovalutazione di cui sopra al capitolo 2. Una valutazione di questo tipo consentirà al lettore di avere una visione d'insieme dei vari problemi di natura linguistica riscontrabili nell'output, prima di passare all'analisi della qualità traduttiva tramite i risultati dei QA checker, così da avere una conoscenza preliminare delle problematiche evidenziate da questi. I parametri che verranno utilizzati a tale scopo saranno la chiarezza del testo tradotto, ovvero il grado di intelligibilità secondo cui il risultato della machine translation risulta comprensibile prima ancora del post-editing, e l'accuratezza, ovvero la corretta trasposizione di dati e informazioni contenuti nel testo source, con particolare attenzioni quindi all'omissione o aggiunta

⁹ JOIN/2015/35/FINAL

di questi. Come precisato nel capitolo precedente, del resto, questi due parametri giocano, l'uno sull'altro vicendevolmente, un'influenza considerevole, motivo per cui, nel momento in cui si andranno ad analizzare eventuali fenomeni che vanno a incidere sulla comprensibilità del testo, sarà inevitabile fare anche riferimento alle eventuali ripercussioni sul piano del carico informativo del messaggio.

Infine, l'ultimo parametro applicabile sarà lo stile. Dei tre, questo sarà sicuramente e per forza di cose il parametro i cui risultati saranno i più interessanti, dato il modo in cui caratteristiche tipologiche e strutturali marcano in maniera netta il linguaggio comunitario, motivo per cui l'utilizzazione di suddetto parametro è sembrata essere imprescindibile.

In seguito, come accennato in precedenza, sarà opportuno servirsi di diversi QA checker per valutare gli standard qualitativi dei tre diversi testi tradotti automaticamente. La necessità di utilizzare più di un QA checker è data dalle funzionalità e caratteristiche, per ovvi motivi differenti, di ogni tool. Per esempio, se da un lato, tanto Matecat quanto SmartCAT sono dotati di applicazioni in-built per il controllo della qualità, dall'altro, SYSTRANet risulta del tutto sprovvisto di suddetta funzionalità, motivo per cui è risultato necessario servirsi di QA checker esterni da applicare uniformemente al prodotto dei tre tool. I QA checker scelti per tale operazione sono XBench e DQF. DQF è stato scelto per un motivo: sebbene Matecat disponga di un editing log in grado di restituire dati sulla produttività, date le ridotte funzionalità del suo QA checker, il quale offre pressoché nessuna possibilità di personalizzazione, è sembrato interessante servirsi dei risultati sulla qualità di un QA checker terzo, implementabile tramite API in Matecat. Tale QA checker è proprio DQF (Dynamic Quality Framework) sviluppato da TAUS.¹⁰ L'utilizzo di questo QA checker sarà duplice: con lo scopo di fornire dati uniformemente validi su tutti e tre i tool si applicheranno i parametri di DQF ai tre, mentre, durante la fase di post-editing su Matecat, si registreranno ulteriori dati tramite l'attività in background dello stesso DQF, data la summenzionata possibilità di implementarne le funzionalità tramite API key.

Infine, alla luce della componente stilistica fortemente marcata dei testi comunitari, ai fini della conduzione di un'analisi a tutto tondo e ispirandosi alla

¹⁰ www.taus.net

metodologia applicata in diversi studi succitati nel capitolo precedente, sarà opportuno valutare la performance dei tre tool a confronto con una versione dello stesso testo, tradotto senza ricorrere alla machine translation.

3.3.1 LA PROSPETTIVA MACROVALUTATIVA: ANALISI DEI PARAMETRI DI CHIAREZZA, ACCURATEZZA E STILE

Sul piano della chiarezza e comprensibilità dell'output dei tre sistemi è possibile affermare con un certo grado di sicurezza che non si riscontrano problemi tanto grandi da inficiare la comprensione del testo, fatte salve alcune problematiche riscontrabili in alcuni segmenti e in maniera più o meno concorde in tutte e tre le traduzioni restituite dai tool. La porzione di testo che ha certamente creato il maggior numero di problemi in tutti e tre i tool è l'elenco dei punti oggetto dell'accordo sul partenariato e lo sviluppo. In maniera differente l'uno con l'altro, ogni tool ha evidenziato differenti problemi sul piano della comprensibilità. Dei tre, sicuramente il tool a mostrare problemi più gravi è stato SYSTRANet, come è possibile osservare nella figura in basso, secondo cui l'obiettivo delle parti coinvolte nell'accordo è la promozione della criminalità organizzata, del riciclaggio di denaro e della corruzione.

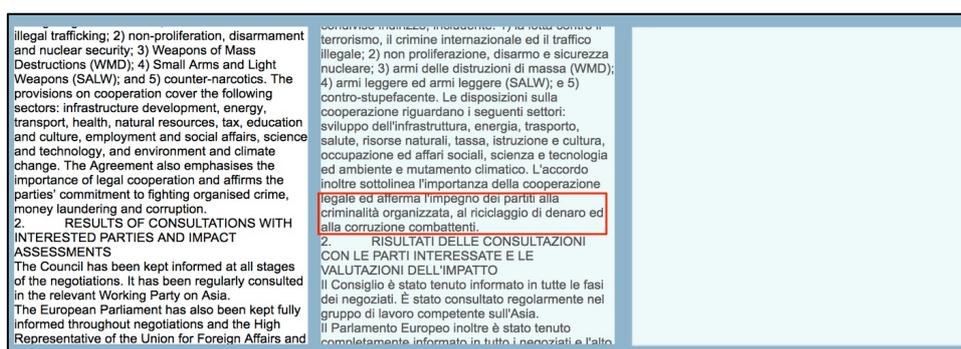


Figura 3.9 Esempio di interpretazione errata del testo source ad opera di SYSTRANet.

Al di là del risvolto più o meno comico che può derivare da una traduzione del genere, è facile vedere come il gerundio “fighting” sia stato interpretato come un sostantivo plurale del tutto slegato dal resto della frase, mentre “commitment” sia

stato attribuito a “organised crime, money laundering and corruption”. D’altro canto, poco prima, come mostra sempre la figura 3.9, SYSTRANet restituisce la seguente traduzione:

- (1) The Agreement builds on the principles of mutual accountability and reiterates the willingness of the parties to address shared concerns, including: (...)

Le configurazioni di accordo per principi di responsabilità reciproca e ripete la volontà dei partiti alle preoccupazioni condivise indirizzo, includente: (...)

(p. 2)

ovvero, una traduzione priva di senso e grammaticalmente errata. Per quanto, come si vedrà più avanti, questo sia stato uno dei passaggi che ha creato il maggior numero di problemi in tutti e tre i casi, SYSTRANet commette anche errori ben più gravi e imperdonabili, come nel caso in cui l’inglese “decade” viene tradotto con “decade”, espressione italiana che indica un lasso di tempo di dieci giorni, quando invece la traduzione corretta è “decennio”. In un’altra occasione, invece, per quanto il periodo elaborato sia lungo e complesso, SYSTRANet restituisce frasi in cui sono presenti vocaboli addirittura non appartenenti alla lingua italiana, come nei due esempi seguenti:

- (2) The analysis further shows that neither of these elements can be regarded as ancillary to any of the other mentioned elements (...)

L’analisi avanza indica che nessuno di questi elementi possono essere considerati come l’accessorio a c’è ne dell’altro ha citato gli elementi (...)

(p. 3)

- (3) The legal basis referred to above confer power on the EU to enter into agreements with third parties (...)

Le basi giuridiche si sono riferite a sopra confer potere sull'UE di prendparteere agli accordi con i terzi (...)

(*ibid.*)

Generalmente, comunque, SYSTRANet dimostra di avere problemi legati all'intelligibilità del testo in quei casi in cui la traduzione suggerita è frutto di un'interpretazione prettamente letterale delle parole del testo source, dove quindi "settled case-law" diviene "giurisprudenza sistemata" mentre "the legal basis must rest on objective factors" diviene "la base giuridica deve riposare sui fattori obiettivi".

Un ulteriore passaggio che, all'interno del testo, ha causato problemi non indifferenti a tutti e tre i tool si presenta come la porzione di testo la cui costruzione è la più articolata o comunque quella che presenta scelte lessicali non del tutto facili da analizzare e ritradurre per la machine translation. Il periodo in questione è il seguente:

A detailed analysis of the aim and content of the Agreement indicates that some of the provisions of the Agreement fall within the scope of the common foreign and security policy, and some other elements fall under the trade policy, as well as under the development cooperation policy of the European Union.

(*ibid.*)

Ciò che è curioso osservare è come il verbo "fall" nella prima occorrenza crei non pochi problemi di resa, mentre, sebbene l'utilizzo sia lo stesso, nella sua seconda occorrenza, questo viene reso correttamente secondo il significato attribuitogli nel dato contesto. Tale fenomeno è osservabile nell'output di tutti e tre i tool, come si evince dai seguenti esempi, presi nell'ordine da Matecat, SmartCAT e SYSTRANet:

- (4) Un'analisi dettagliata dello scopo e del contenuto dell'accordo indica che alcune disposizioni della caduta accordo nell'ambito della politica estera e di sicurezza comune, e alcuni altri elementi rientrano nella politica commerciale, così come nel quadro della politica di cooperazione allo sviluppo dell'Unione europea.
- (5) Un'analisi dettagliata dell'obiettivo e il contenuto dell'accordo indica che alcune delle disposizioni della caduta accordo nell'ambito della politica estera e politica di sicurezza e alcuni altri elementi rientrano nell'ambito della politica commerciale, così come sotto la politica di cooperazione allo sviluppo dell'Unione europea.
- (6) Un'analisi dettagliata dello scopo e del contenuto dell'accordo indica che alcune delle disposizioni della caduta di accordo nell'ambito della politica estera e di sicurezza comune ed alcuni altri elementi rientrano nella politica commerciale come pure nell'ambito della politica di cooperazione allo sviluppo dell'Unione Europea.

Infine, per la traduzione di “Other legal considerations” SYSTRANet propone la totalmente errata traduzione “Altre considerazioni di ordine morale” (p. 4), in una posizione sensibile come può essere quella del titolo di sottoparagrafo e quindi destinata a passare tutt’altro che inosservata. Gli altri due sistemi, d’altro canto, suggeriscono una traduzione adeguata, ma solo quella di Matecat può essere accettabile, dal momento che la traduzione, “Altre considerazioni legali”, evidente calco dall’inglese, proposta da SmartCAT non è accettabile sul piano stilistico, nonché della terminologia consolidata.

Leggermente superiore risulta essere la performance di Matecat, per quanto concerne la chiarezza del suo output. Al di là dei casi succitati, in cui anche Matecat mostra alcune pecche, c’è da notare una tendenza non del tutto infrequente a lasciare non tradotte determinate espressioni, dimostrando, tra i tre, di essere il tool con il più alto numero di parole non tradotte e direttamente trasferite nel testo target. Tale tendenza, come si vedrà in seguito, incide negativamente anche sul parametro dell’accuratezza, dal momento che qualsiasi informazione non tradotta può essere

considerata come un'omissione di parte del significato globale del messaggio nel testo target. È questo il caso dell'espressione "Weapons of Mass Destructions (WMD)". Come si evince in figura 3.10, Matecat traduce solo parzialmente tale espressione e ne restituisce infatti il resto in inglese.

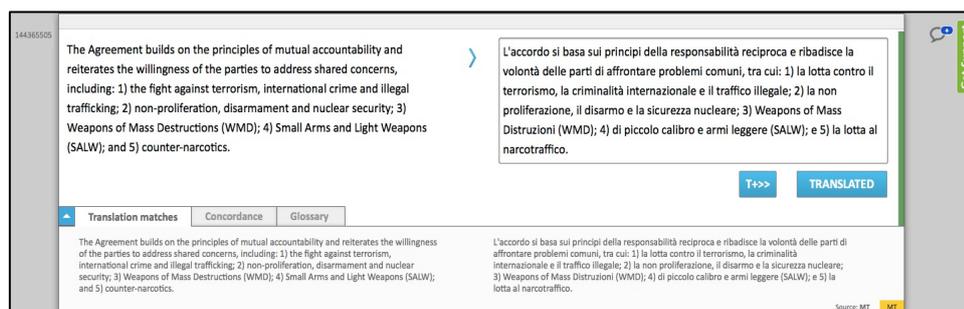


Figura 3.10 Esempio di testo source non tradotto presente nel target in Matecat. Come si può vedere, "Weapons of Mass Destructions (WMD)" è stato tradotto solo parzialmente e l'acronimo è stato lasciato intatto.

Per quanto riguarda l'acronimo, tutti e tre i tool restituiscono la stessa soluzione, ovvero WMD, sebbene, tramite una banale ricerca in rete limitata per mezzo di operatori booleani ai siti aventi per dominio europa.eu, il numero di occorrenze di suddetto acronimo non sia tale da giustificare ed accettarne la presenza in quanto acronimo consolidato nei testi comunitari italiani. D'altro canto, l'acronimo trasposto in italiano ADM restituisce un numero leggermente superiore di occorrenze. Quello degli acronimi, tuttavia, è un discorso più complicato, il quale merita la giusta osservazione, alla luce del diverso comportamento mostrato dai tre tool. Per esempio, se da un lato, tanto Matecat quanto SmartCAT restituiscono il corrispettivo acronimo italiano del Servizio europeo per l'azione esterna, SEAE, dal corrispettivo EEAS (European External Action Service), SYSTRANet riporta nel target l'acronimo così come figura nel testo source. Caso differente, invece, è quello dell'acronimo di "Cooperation Agreement on Partnership and Development (CAPD). In questo caso, SmartCAT e SYSTRANet non traducono l'acronimo, riportandone la versione inglese anche nel target, mentre Matecat riporta l'acronimo (DPCA). Ciò è probabilmente dovuto al fatto che gli acronimi CAPD e DPCA sono effettivamente attestati nell'italiano comunitario, se non fosse che in questa accezione rispondono all'acronimo di Dialisi peritoneale continua ambulatoriale, come si evince da una ricerca in rete limitata ai siti del diritto comunitario dell'UE.

Contrariamente alla summenzionata tendenza di Matecat a riportare parole non tradotte nel testo target, questi risulta essere l'unico tool a proporre una traduzione per l'espressione "multipillar approach", ovvero "approccio multipilastro" (p. 3). Tuttavia, è necessario precisare che la stessa espressione "multipillar approach" è assente in banche dati terminologiche come IATE e molto poco frequente nei portali del diritto comunitario, lì leggermente più frequente nella grafia "multi-pillar approach". Infine, ad inficiare sulla comprensibilità del testo influisce l'errata traduzione del seguente punto

- (7) The Council General Secretariat shall establish the instrument of full powers to sign the Agreement, subject to its conclusion, for the person(s) indicated by the negotiators of the Agreement.

Il segretariato generale del Consiglio definisce lo strumento dei pieni poteri a firmare l'accordo, con riserva della sua conclusione, per la persona (s) indicato dai negoziatori dell'accordo.

È chiaro come Matecat, al contrario di SmartCAT, ma in maniera simile a SYSTRANet, non sia riuscito a cogliere l'espressione della singolarità/pluralità mediante la s parentetica.

SmartCAT, dal canto suo, sebbene, come si vedrà più avanti, risulti essere il tool a restituire le prestazioni migliori in termini di resa stilistica appropriata al contesto comunicativo in questione, dimostra, talvolta, di cadere in traduzioni estremamente letterali, basate su calchi, e quindi errate. Sono questi i casi, per esempio, della resa di "education" con "educazione" in luogo di "istruzione" o di "memorandum esplicativo" in luogo del termine consolidato "relazione". Un ulteriore caso, dato dalla non comprensione della categoria grammaticale della parole in questione, è quello mostrato in figura 3.11.

54	Article 37 TEU grants the EU the power to conclude agreements with third parties in the area of common foreign and security policy.	Articolo 37 borse TEU all'UE il potere di concludere accordi con terzi nella zona di politica estera e di sicurezza comune.	✓	⚠	MT
55	Article 207 TFEU grants the EU the power to conclude agreements in the area of common commercial policy.	Articolo 207 sovvenzioni del TFUE l'Unione europea il potere di concludere accordi nel settore della politica commerciale comune.	✓	⚠	MT
56	Article 3(1) TFEU gives the EU exclusive competence in this area.	Articolo 3, paragrafo 1 del TFUE conferisce all'UE la competenza esclusiva in questo settore.	✓	⚠	MT
57	Article 209 TFEU grants the EU the power to conclude agreements with non-EU countries in the area of development cooperation policy.	Articolo 209 sovvenzioni del TFUE l'Unione europea il potere di concludere accordi con paesi extracomunitari nell'area della politica di cooperazione allo sviluppo.	✓	⚠	MT

Figura 3.11 Traduzione automatica di SmartCAT dei segmenti 54-57.

In questo caso appare chiaro come il sistema MT di SmartCAT abbia interpretato il verbo “to grant” alla terza persona singolare come il sostantivo plurale “grants”, traducendolo prima con borse e successivamente con sovvenzioni, mentre il verbo “to give” è stato riconosciuto quale terza persona singolare del suddetto verbo e tradotto con “conferisce”, un verbo che, limitatamente a registro e tono, risulta perfetto nel contesto comunicativo in questione. Un verbo che invece, per quanto accettabile sul piano pragmatico, non trova un gran numero di occorrenze nei documenti comunitari è “parafare”, individuabile, coniugato al participio passato, nell’output di SmartCAT, quale traduzione di “initialled”, per il quale Matecat e SYSTRANet restituiscono il più consolidato “siglato”. Inoltre, tale suggerimento evidenzia anche incoerenza terminologica, dal momento che invece in altre occasioni, “initialled” è stato reso da SmartCAT con il termine appropriato.

D’altro canto, SmartCAT risulta l’unico tool a non essere in grado di restituire una traduzione accettabile del titolo stesso del documento, ovvero la dicitura dell’accordo di cooperazione sul partenariato e lo sviluppo, presentando una situazione di questo tipo:

(8) (...) on the signing, on behalf of the European Union, of the Cooperation Agreement on Partnership and Development between the European Union and the Islamic Republic of Afghanistan

(...) relativa alla firma, a nome dell’Unione europea, della onPartnership di accordo di cooperazione e sviluppo tra l’Unione europea e la Repubblica islamica di Afghanistan

(p. 1)

Passando a ispezionare quegli elementi che invece contribuiscono all'accuratezza dell'output della MT dei tre sistemi, è possibile affermare che Matecat è il tool a riportare il maggior numero di *inaccuracies*. Ad esempio, questi riporta diverse omissioni di considerevole importanza, come nel caso a seguire:

- (9) Article 209 TFEU grants the EU the power to conclude agreements with non-EU countries in the area of development cooperation policy.

L'articolo 209 del TFUE conferisce all'Unione europea il potere di concludere accordi con paesi terzi nel settore della politica di cooperazione allo sviluppo.

(p. 3)

In questa occasione, la machine translation ha ommesso la traduzione dell'espressione "non-EU" presente nel testo inglese ma non in quello italiano, dove sarebbe dovuta essere resa con "extracomunitari", così come appare tradotta infatti dagli altri due sistemi.

Oltre al precedente caso di omissione, nell'output di Matecat sono presenti ulteriori errori riguardanti la trasposizione di informazioni dal testo source al testo target. Per comprendere meglio il fenomeno, si osservi la figura 3.11.

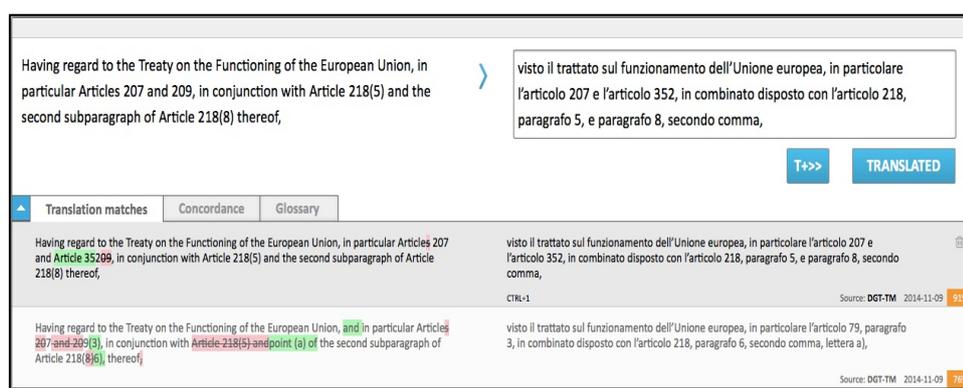


Figura 3.11 Esempio di errata trasposizione di valori numerici nella vista Editor di Matecat. In questa occasione, non vi è concordanza nei valori numerici tra testo source e testo target.

In questa occasione, tuttavia, non avendo la possibilità di agire direttamente sul funzionamento di Matecat e scegliere da dove ricevere i suggerimenti, siano essi provenienti da memorie di traduzione o MT, questo non restituisce alcun suggerimento dalla machine translation, ma piuttosto quello delle memorie di traduzione rese pubbliche dalla DGT. Dal momento che, tra le opzioni del tool, non vi è la possibilità di escludere le memorie di traduzione e limitarsi all'utilizzo della machine translation, nel momento in cui si ordina al tool di pre-tradurre il documento, questi, in mancanza di traduzione automatica, restituisce la traduzione con la percentuale maggiore tra le memorie di traduzione disponibili. Pertanto, quello qui descritto è un errore dovuto a una personalizzazione non del tutto completa delle fonti da cui è possibile visualizzare i segmenti. Presumibilmente, un maggiore grado di personalizzazione è consentito nella versione open-source di Matecat, non presa in esame in questa sede.

D'altro canto, nell'output di Matecat occorrono anche casi di conflitti traduttivi e mancata coerenza terminologica tra segmenti. È questo il caso del seguente esempio.

(10)The Agreement sets up an institutional framework composed of the Joint Committee (see Title VIII, Article 49 *Institutional framework*). The Joint Committee may set up special committees or working groups to assist it with the performance of its tasks.

L'accordo istituisce un quadro istituzionale composto dal comitato misto (vedi titolo VIII, l'articolo 49 *Quadro istituzionale*). La commissione mista può istituire commissioni speciali o gruppi di lavoro per assisterlo con le prestazioni dei suoi compiti.

Nei due segmenti, il sistema traduce correttamente, nella sua prima occorrenza, "Joint Committee" con comitato misto, mentre nella seconda, suggerisce l'errata traduzione "commissione mista", presentando all'utente un caso di incoerenza terminologica. Ciò è del resto dovuto, come è emerso dalla prova pratica, dal fatto che nella seconda occorrenza il termine è preceduto dall'articolo maiuscolo "The".

Per quanto concerne il livello di accuratezza di SmartCAT e SYSTRANet, invece, le prestazioni si attestano su un livello superiore a Matecat, dal momento che non si

riscontrano omissioni o conflitti traduttivi, fatto salvo ad esempio l'alternarsi in SmartCAT, tra diversi segmenti, dell'acronimo stante per il Trattato sull'Unione europea (TUE), a volte resa nella forma italiana, a volte in quella inglese (TEU), in casi per altro in cui l'acronimo è preceduto dagli stessi identici elementi e non vi è punteggiatura particolare.

Limitatamente alla resa stilistica, la prestazione di SmartCAT risulta essere migliore rispetto a quelle di Matecat e SYSTRANet. Fatte salve determinate problematicità strutturali, globalmente, registro e tono risultano essere più in linea con il contesto comunitario. La resa stilistica, infatti, risulta consona a quella richiesta dai testi giuridici, siano essi comunitari o meno. Più in generale, nonostante l'intrinseca impossibilità della machine translation di riprodurre nel suo output le caratteristiche tipologiche e stilistiche dell'euroletto, alla luce delle suddette, SmartCAT risulta comunque il tool in grado di restituire frasi la cui struttura risulta essere la più vicina possibile a quella dell'euroletto, sebbene neanche questi, in determinati contesti, riesca a restituire la corretta traduzione di termini consolidatisi nel lessico comunitario in seguito a un innalzamento stilistico. È questo il caso, ad esempio, di "minori", corretta traduzione di "children", vocabolo italiano arrivato a soppiantare il più comune "bambini".

Per quanto concerne la numerazione e classificazione di atti e articoli in paragrafi, sotto paragrafi, comma e punti, la prestazione di SmartCAT risulta essere sopra la media, se paragonata a quelle di Matecat e SYSTRANet, i quali, in maniera non costante, riportano una dicitura errata dei suddetti elementi di carattere esclusivamente giuridico. Limitatamente alle convenzioni redazionali comunitarie, e in una certa misura specifiche dell'italiano quale lingua di arrivo del documento, la resa dei nomi di istituzioni e cariche in seno all'UE evidenzia prestazioni mediocri da parte di tutti e tre i tool, i quali alternano l'uso di minuscole e maiuscole in maniera incoerente nell'arco dell'intero documento, oltre a un uso scorretto della punteggiatura per segnalare l'inizio e la fine del discorso riportato da altri testi comunitari.

Una menzione particolare meritano le convenzioni specifiche in merito alla nomenclatura dei documenti della Commissione. Nella versione inglese, a seguire il numero dell'atto, è presente la dicitura "final", resa in maniera differente dai tool in

questione. In linea teorica, nessuna delle soluzioni può considerarsi del tutto sbagliata, poiché per un dato periodo, prima una, poi l'altra, tutte queste diciture sono state adottate nei documenti comunitari. Nella fattispecie, la dicitura "final" è e deve essere adottata nei soli documenti redatti a partire dal 27 gennaio 2012; la dicitura "def." doveva essere adottata nei documenti antecedenti l'anno 2000; infine, la dicitura "definitivo" doveva essere adottata nei documenti redatti tra l'anno 2000 e l'anno 2012 (UE: 2011). Per tale motivo, né la soluzione di SmartCAT né quella di Matecat può essere considerata del tutto corretta. Sfortunatamente, per motivi che si andranno a descrivere a seguire, non è stato possibile analizzare le prestazioni di SYSTRANet sul suddetto punto del testo.

3.3.2 IMPLEMENTAZIONE DEI PROCESSI DI REVISIONE E QUALITY ASSURANCE: XBENCH E DQF

Prima di esaminare nel dettaglio i risultati restituiti dai QA checker presi a modello è opportuno fare una premessa sulle prestazioni dei tre tool in funzione del formato dei documenti utilizzati per questo documento, dal momento che ogni tool ha dimostrato tendenze diverse i cui effetti, come si vedrà, si ripercuotono sulla valutazione globale della qualità dell'output della MT. La grandissima parte dei documenti disponibili su Eur-lex, il portale del diritto comunitario, sono disponibili in almeno tre formati: PDF, DOC e HTML. I documenti pubblicati sulla Gazzetta ufficiale sono disponibili anche nella versione digitale della Gazzetta stessa. Per questo studio si è scelto di utilizzare il formato PDF per un motivo semplice: nelle versioni DOC, contrariamente a quelle PDF, il frontespizio di ogni documento contiene un'immagine di testo in cui sono contenute informazioni rilevanti quali l'organo emanante il documento, il numero del documento e il titolo. Un'immagine di testo di questo tipo non sarebbe stata tradotta dai tool, provocando così una perdita di dati e informazioni di grande importanza ai fini del presente studio. Dopo aver optato per questo formato, tuttavia, si è resa necessaria una considerazione ulteriore, data da due considerevoli limiti imposti da SYSTRANet: la funzione di traduzione di file di suddetto tool non consente l'elaborazione di file PDF e, qualora ci si serva invece dell'interfaccia per la traduzione di testo, questa ha un limite massimo di

caratteri, equivalente a 10001 caratteri. Per i motivi summenzionati, si è scelto di utilizzare, pertanto, il file DOC del suddetto documento, sebbene questo abbia comportato una perdita considerevole di testo. Appare inevitabile, quindi, come problematicità di questo tipo influiscano negativamente sulle prestazioni dei tre tool, senza contare, inoltre, che, sebbene Matecat e SmartCAT siano in grado di elaborare file PDF, la segmentazione del documento non è sempre stata precisa e questo ha comportato inevitabili carenze sul piano della resa traduttiva, considerato che, come detto in precedenza, seguire una segmentazione standard e corretta è di fondamentale importanza per il corretto funzionamento di CAT tool e MT. Questo aspetto risulterà importantissimo quando si discuteranno i risultati e si andranno a trarre le conclusioni di questo studio.

Ciò detto, si passerà ad analizzare il funzionamento dei tre tool associati a Xbench. Anche in questo caso, i tre tool evidenziano seri problemi legati all'intrinseca incompatibilità di questi con applicazioni che non sono state progettate specificatamente per questi. Xbench consente di lavorare con una grandissima varietà di formati, due tra i quali sono XLIFF e TMX. Mentre SYSTRANet non consente l'esportazione di nessuno dei due formati, Matecat e SmartCAT consentono la sola esportazione di memorie TMX. Tuttavia, i file TMX esportati dai due tool non sono risultati essere compatibili con Xbench, motivo per cui si è reso necessario un procedimento intermedio, ovvero processare i file TMX con un CAT tool come Trados Studio e impostare un progetto con le suddette memorie al fine di ottenere gli equivalenti file XLIFF.¹¹ I file ottenuti dai due tool, tuttavia conservano, in linguaggio HTML, tutti i tag necessari alla formattazione del testo così come è visibile nella versione pubblicata, rendendo la revisione difficoltosa e comunque non compatibile con Xbench. A fronte di queste difficoltà, in parte superabili, ma richiedenti un quantitativo di tempo extra assolutamente non compatibile con gli standard produttivi richiesti per la machine translation, si è deciso di non approfondire oltre l'interazione dei tool con Xbench, dal momento che in un ipotetico contesto lavorativo questa soluzione sarebbe stata immediatamente scartata.

¹¹ Essendo sprovvisto di una funzione che consente l'esportazione di file TMX, per quanto riguarda SYSTRANet, è bastato allineare i due testi paralleli, ovvero la versione inglese e il testo tradotto automaticamente, per ottenere un file TMX.

È inevitabile, pertanto, come ciò abbia decretato un'incompatibilità di fondo tra i tool presi in esame e il programma di revisione.

Come detto in precedenza, al contrario di SYSTRANet, Matecat e SmartCAT sono provvisti di QA checker e funzionalità di revisione propri. Al fine di ottenere dati obiettivi e dettagliati sulla qualità dell'output della MT, si è quindi reso opportuno comparare i risultati di suddetti QA checker in-built con un QA checker esterno, applicabile a tutti e tre i tool, ovvero DQF. La valutazione della qualità tramite DQF viene condotta in maniera estremamente semplice: al valutatore viene richiesto di esprimere, per ogni segmento, un giudizio su due parametri, "Fluency" e "Adequacy", e di individuare e classificare gli errori riscontrati secondo cinque tipologie che si andranno a elencare e descrivere a seguire. Le linee guida della TAUS applicate a DQF definiscono il parametro "Fluency" come il grado secondo cui la traduzione è comprensibile, facilmente leggibile e interpretabile da un lettore nativo della lingua target, nonché grammaticalmente e lessicalmente accurata. Tale valutazione può essere condotta assegnando a ogni segmento un punteggio che va da 4 a 1, dove 4 corrisponderà a una traduzione ineccepibile dal punto di vista della scorrevolezza e comprensibilità, mentre 1 corrisponderà a una traduzione del tutto incomprensibile. Per quanto riguarda, invece, il parametro "Adequacy", questo è definito come il quantitativo di informazioni che sono presenti tanto nel testo source quanto nel testo target, senza aggiunte né omissioni. Anche in questo parametro, la valutazione si snoda lungo una scala da 4 a 1, dove a nessuna perdita di informazioni corrisponderà il punteggio massimo, mentre 1 corrisponderà a un segmento in cui nessuna delle informazioni del testo source è stata trasferita nel testo target. Questa definizione di adeguatezza del testo tradotto risulta essere equivalente al parametro dell'accuratezza adottato nel paragrafo precedente. È opportuno precisare, infatti, che quello dell'adeguatezza è un criterio che incide ed è influenzato da un più ampio spettro di fattori, differenti a seconda dei destinatari e degli utilizzatori ultimi della traduzione stessa, concetti dinamici e olistici che variano a seconda delle circostanze. Per tale motivo, pertanto, sebbene in questa accezione, per traduzione adeguata si intenderà quella che non tralascia alcun elemento del contenuto referenziale, è tuttavia opportuno ricordare che, ad esempio, per il traduttore una traduzione adeguata è una traduzione in cui il suo dispendio di forze e tempo è proporzionale al

successo comunicativo e agli standard di qualità prefissati, mentre per il committente l'adeguatezza è più legata a fattori di ordine economico (Scarpa, 2008: 212-213).

L'individuazione e categorizzazione degli errori, come si è visto, può variare di valutazione in valutazione, seppure siano riscontrabili determinati parametri costanti in ognuna di esse. Un approccio valutativo di questo tipo risulta essere una metodologia di valutazione standard, applicata in maniera consistente nel settore, con le dovute differenziazioni per quanto concerne, come si è detto, categorie, penalità associate in base all'entità dell'errore e così via. Trattandosi di un processo manuale, conducibile da un individuo con ottima padronanza tanto della lingua source quanto della lingua target, questi si rivela, limitatamente a uno studio comparativo di questo tipo, una metodologia di valutazione con considerevoli vantaggi dal punto di vista del tempo e delle risorse spese. Inoltre, uno scenario in cui tale approccio dimostra la sua efficacia è proprio quello della valutazione dell'output della machine translation, al fine di individuare errori ricorrenti in esso e approntare strategie di miglioramento delle sue prestazioni.¹² Il numero di categorie definibili può variare a seconda della tipologia testuale, dello scopo della traduzione e, ovviamente, del grado di approfondimento e differenziazione degli errori che si vuole ottenere dalla propria valutazione. Ogni categoria, comunque, dovrebbe disporre di una descrizione quanto più precisa possibile degli errori ascrivibili alla suddetta. Delle diverse categorie comuni a bene o male tutte le valutazioni della qualità traduttiva, quella che forse più di tutte necessita di una descrizione chiara e ben definita è quella dello stile, dal momento che, in mancanza di essa, può subentrare un considerevole grado di soggettività che potrebbe invalidare i risultati della valutazione. A tale scopo, può essere di aiuto semplificare la questione a una semplice ma efficace definizione: un errore di stile è un errore che contravviene alla guida di stile consolidata e adottata nella traduzione in oggetto.¹³

DQF consente l'individuazione e categorizzazione degli errori riscontrabili nella traduzione oggetto di valutazione, ma non include la possibilità di applicare penalità diverse a seconda della gravità dell'errore, né comprende le threshold entro cui

¹² Contenuto riadattato da linee guida ed enciclopedia multimediale disponibili sul sito Internet di TAUS, di cui si fornirà il link completo nella bibliografia.

¹³ Contenuto adattato dal video "Error Typology Quality Evaluation", breve estratto dalla conferenza TAUS Quality Evaluation Summit, di cui si fornirà il link completo nella bibliografia.

valutare se la traduzione può passare la valutazione o meno. Per tale motivo, in questa sede ci si limiterà a individuare e categorizzare gli errori commessi nell'output dei tre tool, un compito più funzionale all'obiettivo di individuazione delle problematiche ricorrenti nei tre sistemi. Le tipologie di errori delineate in DQF non differiscono enormemente da quelle più adottate nel mercato e le linee guida di TAUS sono pressoché equivalenti a quelle delle più note società specializzate nella creazione di parametri di valutazione della qualità traduttiva. Nella tabella 3.1 si fornisce una panoramica descrittiva delle cinque tipologie di errori.

Accuratezza	Consiste nell'individuazione di quegli errori dovuti a una errata interpretazione del testo source. Rientrano in questa categoria anche omissioni e aggiunte di informazioni, nonché porzioni di testo non tradotte e trasferite nel testo target.
Lingua	In questa categoria rientrano tutti gli errori di tipo grammaticale e sintattico, nonché ortografico. Anche gli errori di punteggiatura vengono inquadrati in questa categoria.
Terminologia	Non rispettare la terminologia consolidata della tipologia testuale oggetto della traduzione comporta la presenza di errori terminologici ascrivibili a questa categoria. L'incoerenza terminologica all'interno dello stesso testo è altresì considerata un errore di questo tipo.
Stile	Non rispettare le guide di stile della società o del cliente, nonché quelle proprie della tipologia testuale oggetto della traduzione, comporta la presenza di errori stilistici. Inoltre, sono considerati errori stilistici anche una traduzione letterale, non fluida e che fa un uso non idiomatico della lingua di arrivo.
Convenzioni specifiche della lingua target	In questa categoria rientrano tutti gli errori dovuti alla non osservanza di regole redazionali convenzionali della lingua target, ovvero date, unità di misura, elementi culturo-specifici e pragmatici.

Tabella 3.1 Categorizzazione degli errori secondo le linee guida adottate da TAUS. Ad ogni categoria corrispondono diversi errori, il cui peso nella valutazione della qualità globale del testo tradotto è lo stesso, venendo questi conteggiati di volta in volta per ogni segmento.

Preliminarmente all'analisi dei dati sulla qualità traduttiva, è opportuno soffermarsi per un momento a confrontare i suddetti parametri con i parametri di revisione di Matecat, poiché dal confronto emergono dati interessanti attraverso cui sarà possibile interpretare i risultati restituiti dai due. Innanzitutto, come si è avuto

modo di osservare precedentemente, le tipologie di errori individuabili e categorizzabili in Matecat non lasciano ampio spazio alla differenziazione, constando di cinque parametri in cui sono raggruppate diverse tipologie di errore che forse meriterebbero una categoria a parte, o quantomeno una descrizione più dettagliata delle categorie a cui ascrivere suddetti errori. Inoltre, a differenza dei parametri di DQF, non viene fatta alcuna menzione circa quegli errori dovuti alla non osservanza delle convenzioni specifiche della lingua di arrivo. Al suo posto, trova invece spazio la categoria deputata alla segnalazione di errori nei tag, gli unici errori, tra l'altro, a venire segnalati automaticamente come critici dal QA checker in-built di Matecat. Per quanto concerne la differenziazione della gravità dell'errore, tuttavia, l'apparato di revisione consente di specificare tra errori critici e non, sebbene questi poi, a differenza dei tag mismatch, non consistano in un ostacolo al salvataggio dell'eventuale traduzione. La differenza sostanziale tra Matecat e DQF è tuttavia un'altra: mentre DQF consente la quantificazione, segmento per segmento, degli errori, specificando la relativa categoria di appartenenza, Matecat consente solo di segnalare che, in un dato segmento, si riscontrano determinate problematiche, siano esse per esempio stilistiche o grammaticali, senza potere tuttavia individuare e isolare eventuali molteplici errori all'interno dello stesso segmento. Per tale motivo appare ovvio come il quantitativo di errori riscontrabili su DQF sia nettamente superiore a quello di Matecat e quindi più incline a restituire un'immagine fedele della qualità dell'output del tool. A seguire, il confronto visivo tra i due quality report, per quanto concerne la tipologia di errori riscontrati.

Job 247495 - Quality Report <small>(en-US > it-IT)</small>		Overall quality Fail <small>Details</small>
Type of Issue	Issues Found / Max tolerated Issues <small>for 1,188 words 1 Enhancement = 0.03 Error</small>	Score
Tag issues (mismatches, whitespaces)	6 / 2.4	Fail
Translation errors (mistranslation, additions/omissions)	21 / 2.4	Fail
Terminology and translation consistency	3 / 3.6	Poor
Language quality (grammar, punctuation, spelling)	21 / 3.6	Fail
Style (readability, consistent style and tone)	12.24 / 5.9	Fail

a) Quality report di Matecat

English (United States) > Italian	
MT Engine: MyMemory	
Number of Segments: 102	
Number of Words: 1682	
Error Type	Number of Errors
Evaluators:	Jacopo
Accuracy (Info)	34
Language (Info)	44
Terminology (Info)	5
Style (Info)	30
Country standards (Info)	2
Total	115

b) Quality report di DQF

Figura 3.12 Confronto tra i Quality report di Matecat e DQF. Il report a) restituisce il numero di errori contro il numero di errori tollerati da Matecat per ogni categoria in base al numero di parole. Il report b) restituisce il numero di errori per ogni categoria e la somma di questi.

Il parametro che, nei due report, restituisce la discrepanza maggiore risulta essere quello dello stile, alla luce della differenza sostanziale illustrata sopra. L'impossibilità imposta da Matecat di quantificare gli errori stilistici che, frequentemente, ricorrono in più di un caso per segmento ha necessariamente dettato un punteggio inferiore, ma comunque insufficiente, rispetto a quello di DQF. D'altro canto, la categoria della terminologia ha restituito risultati più o meno equivalenti, dal momento che, sommariamente, gli errori terminologici riscontrati nell'output di Matecat si sono limitati ai casi illustrati in precedenza, dettando quindi una maggiore facilità e precisione nell'individuazione di questi.

Passando in rassegna gli errori individuati da DQF nell'output di SmartCAT e confrontandoli con gli errori individuati da quest'ultimo emergono risultati differenti da quelli appena riportati, come si evince in figura 3.13.



a) Quality report di SmartCAT

English (United States) > Italian

MT Engine: Bing Translator
 Number of Segments: 112
 Number of Words: 1633

Error Type	Number of Errors
Evaluators:	User#1
Accuracy (Info)	33
Language (Info)	39
Terminology (Info)	6
Style (Info)	27
Country standards (Info)	1
Total	106

b) Quality report di DQF

Figura 3.13 Confronto tra i Quality report di SmartCAT e DQF. Il report a) indica il numero di errori totali per categoria individuati da SmartCAT.

Il QA checker di SmartCAT consente di impostare diversi parametri secondo le proprie preferenze, sebbene tuttavia non vi sia alcuna possibilità di valutare, nella fase di revisione, parametri fondamentali quali la resa stilistica, né tantomeno la grammatica e la sintassi del testo tradotto, nonché le scelte traduttive stesse. Dal confronto tra i due report, tuttavia, emerge, primo su tutti, un dato: il report di SmartCAT segnala ben più errori di DQF per quanto concerne la terminologia, nonché un numero piuttosto consistente per quanto riguarda l'aspetto più inerente la grammatica. Ciò è dovuto alla tendenza di SmartCAT a restituire diversi falsi positivi, ovvero rese traduttive corrette della machine translation ritenute tuttavia errori dal QA checker. Questo è ad esempio il caso dei numerosissimi errori raggruppati nella categoria "Punctuation" dove trovano spazio tutti quegli elementi, quali acronimi o formule fisse dei documenti comunitari, che per ragioni redazionali sono scritti interamente in maiuscolo e quindi considerati come errori dal tool.

Analogamente, nella categoria “Other” rientrano errori come l’esatta coincidenza tra source e target, dovuta in questo caso all’intraducibilità, ad esempio, del numero identificativo della proposta di decisione. Infine, nella categoria “Terminology” diversi errori sono dovuti al fatto che, specificando un glossario di progetto, il tool è portato ad individuare come errori tutti quei casi in cui, a prescindere dalla categoria grammaticale e dalla conseguente concordanza, il termine è reso in maniera diversa da quella specificata nel glossario. È questo ad esempio il caso di “ruling”, equivalente di “sentenza” nel glossario ma considerato errore nel momento in cui, ad esempio, “rules” viene tradotto con “norme”.

Per quanto riguarda il numero di errori riscontrati nel confronto tra Matecat e SmartCAT è possibile vedere come la prestazione di SmartCAT sia leggermente superiore a quella di Matecat, soprattutto per quanto concerne la resa stilistica, mentre i punteggi nelle altre categorie si attestano su valori mediamente vicini tra loro.

Limitatamente alla valutazione della qualità dell’output di SYSTRANet, non disponendo di un proprio QA checker, non è stato possibile operare un confronto tra questo e DQF. Ciononostante, la sola valutazione condotta tramite DQF consente di analizzare la performance di SYSTRANet a sé stante e in correlazione con gli altri due tool. A seguire, si fornisce il Quality report di DQF riguardo la valutazione degli errori riscontrabili in SYSTRANet.

English (United States) > Italian	
MT Engine: Systran	
Number of Segments: 79	
Number of Words: 1601	
Error Type	Number of Errors
Evaluators: User#1	
Accuracy (Info)	54
Language (Info)	39
Terminology (Info)	16
Style (Info)	61
Country standards (Info)	1
Total	171

Figura 3.14 Quality report di DQF dell’output di SYSTRANet.

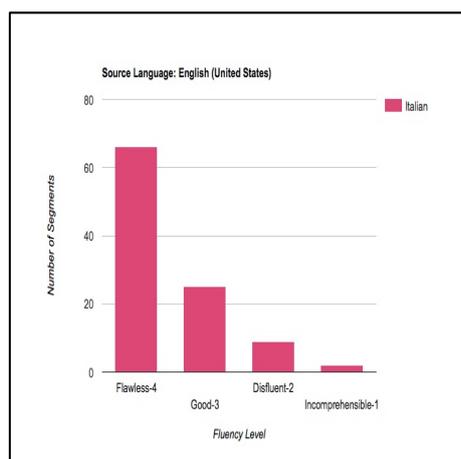
Dal Quality report di DQF emergono in maniera lampante le problematiche riscontrate in precedenza nell’output di SYSTRANet, confermate dal numero

considerevolmente maggiore di errori totali riscontrati. Il parametro a influire in maniera più consistente è ancora una volta lo stile, mentre è importante notare come il numero di errori ascrivibili alla categoria “Language” sia equivalente a SmartCAT e più basso di Matecat, dal momento che, laddove gli altri due tool hanno evidenziato problemi di concordanza nel genere e numero dei sostantivi e un’errata scelta del tempo e modo verbale, SYSTRAnet ha dimostrato una maggiore cura e precisione per questo aspetto.

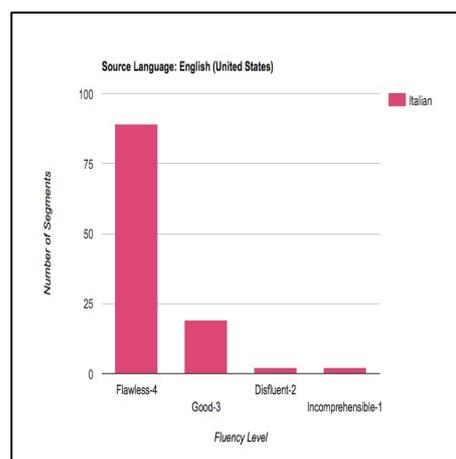
A seguire, adesso, invece si passeranno ad comparare i risultati ottenuti dai tre tool per quanto riguarda i parametri di “Fluency” e “Adequacy” valutati in DQF. Una valutazione di questo tipo è adottabile in quei contesti in cui è necessario accertarsi che il contenuto tradotto comunichi le stesse informazioni del testo originale, in modo corretto e conforme alla lingua di arrivo e in maniera chiara, fluida e scorrevole.¹⁴ Una valutazione allo stato dell’arte di questo tipo deve includere linee guida precise ed estremamente chiare, al fine di ottenere dati rilevanti e coerenti quando applicati su diversi sistemi. Allo stesso modo, anche i criteri tramite i quali valutare una traduzione in funzione di questi due parametri dovrebbero essere chiari e di facile, nonché immediata, comprensione e applicazione.¹⁵ A seguire si forniscono i risultati ottenuti dai tre tool nel parametro “Fluency”.

¹⁴ Cfr. nota 12.

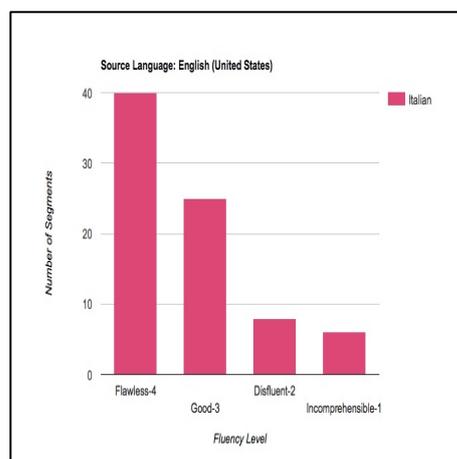
¹⁵ Contenuto adattato dal video “Adequacy, Fluency and/or Readability Evaluation”, breve estratto della conferenza TAUS Quality Evaluation Summit (2013), di cui si fornirà il link in bibliografia.



a) Matecat



b) SmartCAT



c) SYSTRANet

Figura 3.15 Grafici dei punteggi ottenuti dai tre tool nel parametro “Fluency”. Nei tre grafici sull’asse x sono disposti i punteggi attribuibili a ogni segmento nella valutazione della “Fluency” di ognuno di essi, secondo i seguenti valori: Flawless-4, Good-3, Disfluent-2 e Incomprehensible-1. Sull’asse y sono invece disposti il numero totale dei segmenti, diverso per ogni tool, a causa delle diverse regole di segmentazione accennate in precedenza. Tali grafici vengono elaborati e riprodotti da DQF a completamento del task di valutazione.

I dati che emergono dai tre grafici confermano quanto evidenziato precedentemente nel capitolo durante l’analisi dei parametri di macrovalutazione applicati ai tre output, ovvero che il tool a restituire il più alto punteggio limitatamente alla leggibilità, chiarezza e scorrevolezza del testo è SmartCAT, per il quale il numero di segmenti con il punteggio più alto ottenibile supera di poco Matecat e di gran lunga SYSTRANet, il quale d’altro canto si conferma come il tool avente il maggior numero di problemi per quello che concerne tale criterio,

coerentemente alle osservazioni presentate. Infatti, SYSTRANet restituisce il più alto numero di segmenti aventi punteggio 1 e 2, nonché il più basso numero di segmenti giudicati “Flawless”. Il numero di segmenti valutati “Good” (3) invece si attesta pressappoco sullo stesso livello per tutti e tre i tool.

Passando ad analizzare i risultati ottenuti dai tre tool nel parametro “Adequacy”, la situazione che si presenta è quella illustrata in figura 3.16.

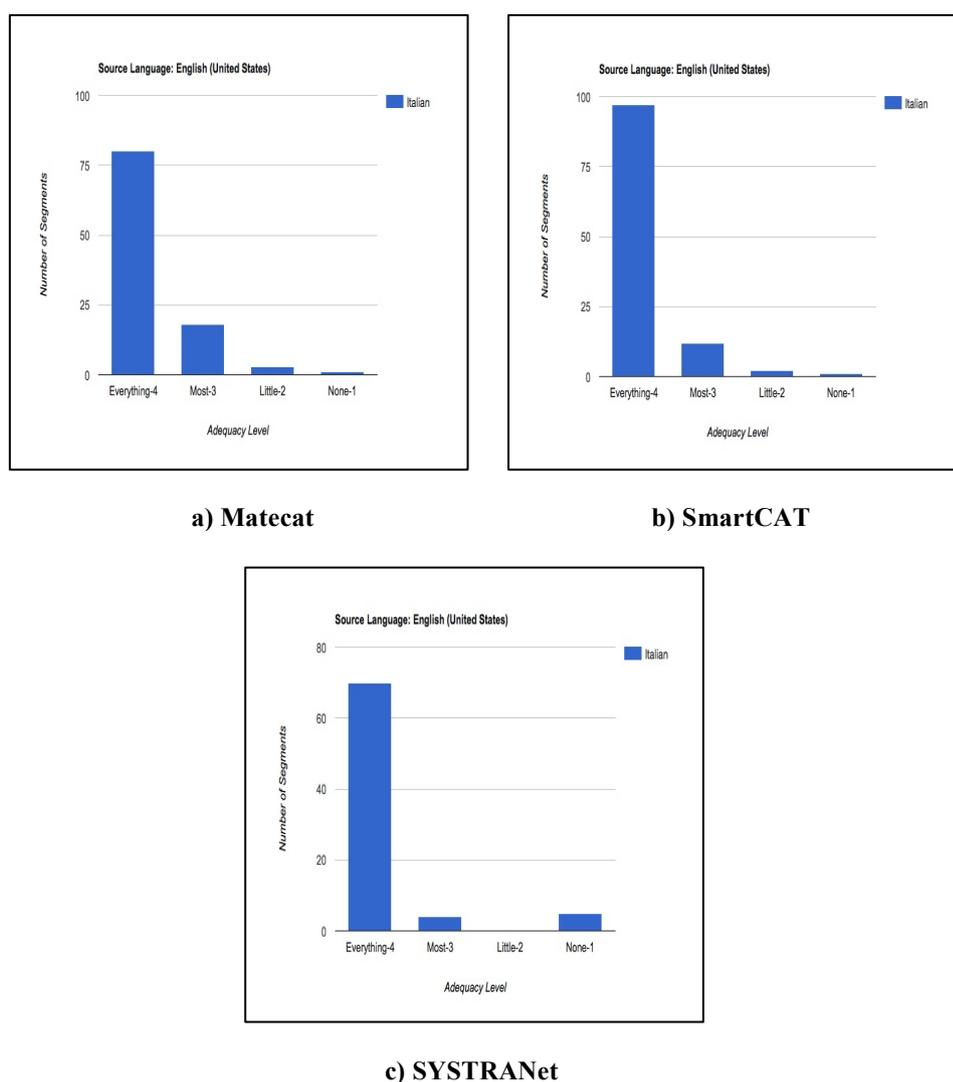


Figura 3.16 Grafici dei punteggi ottenuti dei tre tool nel parametro “Adequacy”. Nei grafici, sull’asse x sono disposti i punteggi ottenuti dai tre tool in base al quantitativo di informazioni del testo source correttamente riportate nel testo target, secondo i seguenti valori: Everything-4, Most-3, Little-2 e None-1. Sull’asse y è presente il numero di segmenti di ogni testo.

Contrariamente al parametro “Fluency”, in questo caso le prestazioni dei tre tool risultano essere generalmente concordi e uniformi tra loro, con leggere oscillazioni

nel numero dei segmenti, che si fanno più importanti solo nel numero di segmenti aventi il punteggio più basso di tutti (None-1), dove SYSTRANet detiene ancora una volta la percentuale maggiore. Ciononostante, a fronte della discrepanza nel numero di segmenti, a parità di numero totale di segmenti, SYSTRANet dimostra di possedere un numero più alto di segmenti aventi punteggio massimo, ovvero “Everything-4”, rispetto a Matecat, coerentemente ai molteplici casi di omissione di informazioni rilevanti, nonché frammenti di segmenti non tradotti, riscontrati in precedenza nell’output del tool avente per sistema MT MyMemory.

3.3.3 TRADUZIONE AUTOMATICA E REFERENCE TRANSLATION A CONFRONTO

Benché dettagliati in quanto in grado di fornire un’immagine completa e veritiera della qualità dell’output dei tre programmi web-based oggetto di questo studio, i risultati di cui sopra mancano di una componente, talvolta inevitabilmente necessaria, talvolta preferibile, ovvero un’analisi comparativa del testo tradotto automaticamente con una versione dello stesso testo tradotto da una persona, senza che questi faccia uso di machine translation. Come ampiamente descritto nel capitolo 1, disporre di una traduzione svolta da un traduttore in carne e ossa, definita reference translation, risulta la componente necessaria alla corretta esecuzione di una valutazione automatica, dal momento che queste valutano quanto una traduzione automatica sia distante da una svolta da un traduttore. Benché raramente utilizzate a sé stanti per valutare adeguatezza dell’informazione tradotta e scorrevolezza del messaggio trasmesso, valutando la precisione del sistema e l’incidenza di segmenti di considerevole lunghezza condivisi tanto dalla MT quanto dalla reference, queste restituiscono feedback per quanto concerne i suddetti parametri. Tuttavia, quando si può disporre di una reference translation, questa può risultare implementabile anche nella valutazione condotta da individui, soprattutto per criteri come l’accuratezza/adequatezza, dal momento che una comparazione dei due prodotti può facilitare l’individuazione di informazioni mancanti, o aggiunte, tanto nell’uno quanto nell’altro.

Data la natura fortemente connotata dei testi comunitari, in virtù dello stile fortemente marcato di questi, applicare una valutazione automatica porterebbe a

ottenere risultati in netto contrasto con quelli illustrati finora, dal momento che sarà quantomeno raro riscontrare in un testo tradotto automaticamente scelte stilistiche ricorrenti quali la tematizzazione del verbo, la diatesi passiva o la nominalizzazione, per citarne alcune. Per tale motivo, è sempre auspicabile prendere in considerazione i risultati di suddette valutazioni in funzione del contenuto del testo che si vuole valutare.¹⁶

Ciò detto, tuttavia, al fine di sostanziare questo studio con ulteriori dati sulle prestazioni della traduzione automatica, traendo ispirazione del meccanismo intrinseco delle *automatic metric*, si procederà ora alla valutazione comparativa dell'output della MT con una reference translation della proposta di decisione presa in esame in questo studio. Per i motivi elencati sopra, tale confronto sarà incentrato sull'analisi dei risultati ottenuti dai due prodotti quando a questi vengono applicate linee guida e metodologia di valutazione di DQF, al fine di dimostrare, in maniera pratica, come alcuni dei fenomeni descritti nel capitolo precedente siano riscontrabili anche nel caso qui esposto. A seguire, si forniscono i risultati riportati da DQF sugli errori individuati nella traduzione di riferimento.

English (United States) > Italian	
MT Engine: Not Specified	
Number of Segments: 87	
Number of Words: 1608	
Error Type	Number of Errors
Evaluators:	
Accuracy (Info)	9
Language (Info)	9
Terminology (Info)	1
Style (Info)	12
Country standards (Info)	0
Total	31

Figura 3.17 Report degli errori riscontrati nella reference translation in DQF

Delle cinque categorie riportate, due sono sicuramente quelle ad attirare l'attenzione massima: Accuratezza e Stile. Con un ragionevole grado di certezza è possibile affermare che dei 9 errori riscontrati nella categoria "Accuracy", salvo un caso, si tratta sempre di omissioni di elementi, più o meno fondamentali alla comprensione globale del messaggio, presenti nel testo source e omessi nel testo

¹⁶ Cfr. nota 12.

target, contrariamente a quanto è stato riscontrabile nella traduzione automatica, per la quale i punteggi generalmente alti in questa categoria erano dovuti principalmente a errori di interpretazione e traduzione del testo source, oltre a elementi non tradotti presenti nel testo source. Questo dato conferma quanto esposto nel capitolo precedente per mezzo di diversi studi presi qui a modello, ovvero che, ponendo maggiore attenzione nella resa stilistica, è probabile riscontrare omissioni di parti del messaggio in un testo tradotto piuttosto che in uno tradotto automaticamente, fatte salve, ovviamente, le eccezioni dibattute in precedenza.

Per quanto concerne lo stile, a questo punto, è possibile osservare senz'altro una maggiore aderenza alle convenzioni stilistiche dei testi comunitari, ovvero a quelle scelte sintattiche e lessicali che solo un traduttore in carne e ossa è in grado di riprodurre in un testo tradotto, poiché facenti capo a conoscenze linguistiche testuali e pragmatiche, impossibili da emulare per la machine translation. Gli errori stilistici segnalati, di conseguenza, sono per lo più dovuti a sporadiche inosservanze delle convenzioni stilistiche comunitarie, quali ad esempio l'uso della maiuscola per indicare enti e cariche in seno all'Unione o espressioni formulaiche cristallizzate e tipiche del linguaggio giuridico, a differenza dell'output della MT, dove sono state riscontrate infrazioni dei manuali di convenzioni redazionali più gravi e, soprattutto, più frequenti.

Infine, gli errori riportati nella categoria "Language" sono imputabili ad errori di punteggiatura riscontrati, ad esempio, negli elenchi puntati e numerati di cui abbonda il testo.

3.4 COME VALUTARE LA PRODUTTIVITÀ DELLA MACHINE TRANSLATION: KEYTRACKER E PRODUCTIVITY EVALUATION (PE)

Dopo aver valutato e comparato la qualità traduttiva dell'output dei tre CAT tool web-based presi qui a modello, al fine di individuare quello la cui prestazione risultasse la migliore, sembra quantomeno opportuno e necessario valutare le potenzialità di questi in funzione degli standard produttivi restituibili da questi a seguito dell'attività di post-editing, con lo scopo di definire quanto questa tecnologia

parzialmente free, web-based e open-source funga da asset per il traduttore e non sia invece un ostacolo all'esecuzione della propria attività.

Per ricavare e presentare dati di questo tipo, ispirandosi a diversi studi presenti nella letteratura in materia, si monitorerà l'attività di post-editing sui tre output, registrando l'attività su tastiera e mouse, comparando le tre prestazioni e correlando quindi tali risultati con la qualità delle tre versioni del testo comunitario. Tale compito verrà eseguito servendosi di Inputlog,¹⁷ un *keytracker* sviluppato dall'Università di Anversa e disponibile per il download gratuito previa registrazione. Infine, come anticipato precedentemente, limitatamente a Matecat ci si servirà delle funzionalità di valutazione della produttività di DQF, implementabili tramite API key in Matecat.

3.4.1 MONITORAGGIO DEL PROCESSO DI POST-EDITING ATTRAVERSO L'UTILIZZO DEL KEYTRACKER INPUTLOG

Progettato e sviluppato da un team di ricercatori dell'università di Anversa, Inputlog consente di tracciare l'attività dell'utente su tastiera e mouse durante una fase di lavorazione su documenti Word, registrando informazioni tra cui durata della fase di lavorazione, numero totale di pressioni sulla tastiera e relativa frequenza, espressa in millisecondi, nonché durate delle pause e punto nel testo in cui queste occorrono. Inoltre, integrando un eye-tracker, è possibile combinare i dati ottenuti dal monitoraggio dell'attività sulle periferiche del computer con i dati ottenuti dal monitoraggio dei movimenti e delle fissazioni oculari. Analizzata la fase di lavoro, Inputlog restituisce un file log in cui sono riportate tutte le informazioni di cui sopra, personalizzabile in base ai dati che si desidera ottenere, escludendo quindi quelli non rilevanti alla propria ricerca, definiti *noise*. Infine, la funzione di riproduzione consente di riprodurre la sessione appena registrata e analizzata, al fine di esaminare eventuali fenomeni ricorrenti e rilevanti nell'attività dell'utente monitorato.

In questa sede, si è utilizzato Inputlog per registrare l'attività di post-editing condotta sui file Word ricavati dal salvataggio dell'output dei tre tool in documenti di testo. Coerentemente alle aspettative nutrite, i dati restituiti da questa operazione

¹⁷ www.inputlog.net

sembrano confermare, al di là di ogni ragionevole dubbio, quanto postulato finora. Qui di seguito, si presenteranno due diverse tipologie di dati raccolti: la prima tipologia è il riepilogo quantitativo dell'attività globale monitorata da Inputlog, mentre la seconda tipologia è un grafico che sussume l'andamento del processo analizzato da Inputlog.

In prima analisi, dal "Summary" è possibile osservare il tempo totale impiegato per posteditare l'output dei tre tool, un dato da correlare alla qualità di questo, poiché maggiori sono le problematicità da risolvere, maggiore sarà il tempo necessario. Questo dato, infatti, conferma quanto detto finora, dal momento che l'output di SYSTRANet (c) risulta essere quello che ha richiesto il maggior tempo, come mostrato in figura 3.18.

Process Time	
General	
Total Process Time (s)	00:33:10
Number of P-Bursts	1,990.190
Mean Process Time P-Bursts (s)	107
Standard Deviation P-Bursts	18.600
Total Pause Time	
Total Pause Time (s)	00:07:20
Number of Pauses	440.220
Mean Pause Time (s)	106
Standard Deviation Pause Time	4.150
Active Writing Time	
Total Writing Time (s)	00:25:49
	1,549.970

a) "Process Time" di Matecat

Process Time	
General	
Total Process Time (s)	00:30:42
Number of P-Bursts	1,842.410
Mean Process Time P-Bursts (s)	104
Standard Deviation P-Bursts	17.720
Total Pause Time	
Total Pause Time (s)	00:06:28
Number of Pauses	388.250
Mean Pause Time (s)	103
Standard Deviation Pause Time	3.770
Active Writing Time	
Total Writing Time (s)	00:24:14
	1,454.160

b) "Process Time" di SmartCAT

Process Time	
General	
Total Process Time (s)	00:46:48
Number of P-Bursts	2,808.920
Mean Process Time P-Bursts (s)	132
Standard Deviation P-Bursts	21.280
Total Pause Time	
Total Pause Time (s)	00:07:30
Number of Pauses	450.740
Mean Pause Time (s)	131
Standard Deviation Pause Time	3.440
Active Writing Time	
Total Writing Time (s)	00:39:18
	2,358.180

c) "Process Time" di SYSTRANet

Figura 3.18 "Process Time" di Matecat, SmartCAT e SYSTRANet Il riepilogo mostra, in minuti e millisecondi, il tempo impiegato totale (General), in pausa (Total Pause Time) e in scrittura (Active Writing Time).

Un altro dato interessante che emerge dalla figura è come il tempo totale sia stato suddiviso tra attività di scrittura (Active Writing Time) e pause (Total Pause Time). Sebbene, come detto in precedenza, correlare pause e qualità produttiva non restituisca dati abbastanza solidi ma piuttosto aleatori perché possano essere utilizzati a proprio supporto, è comunque possibile ipotizzare che, alla luce di una qualità inferiore dell'output, il tempo trascorso in pausa sarà presumibilmente maggiore, poiché è necessario un maggiore sforzo cognitivo per post-editare il testo. Infatti, dei tre tool, quello ad avere registrato il maggiore tempo trascorso in pausa è SYSTRANet (c), seguito con uno scarto di circa 10 secondi da Matecat (a). Soffermandosi ancora un momento sulle pause registrate da Inputlog, inoltre, è possibile ricavare altre informazioni rilevanti da dati di semplice interpretazione, quali il numero totale di pause registrate dal programma, dato che ancora una volta questi pone SYSTRANet (c) in prima posizione, seguito da Matecat (a) e infine da SmartCAT (b), il quale, limitandosi ai dati elencati finora risulta essere il tool dalle prestazioni migliori. Per quanto sia difficile ipotizzare una spiegazione, correlabile con natura e scopo di questo studio, è tuttavia opportuno segnalare che la durata media delle pause di Matecat, poco più di 4 secondi, risulta essere la maggiore tra i tre tool. È forse possibile ipotizzare che queste siano occorse in punti di più difficile risoluzione, con un'incidenza tuttavia minore rispetto a SYSTRANet.

Un altro dato di semplice utilizzo e analisi ma cionondimeno recante informazioni preziose è, semplicemente, il numero di parole e caratteri digitati registrati da Inputlog, dato osservabile in figura 3.19.

Process Information	
Characters	
Total Typed in Main Document	11704
Total copied	11009
Total typed (incl.spaces)	695
Per minute (incl. spaces)	20.953
Total typed (excl.spaces)	608
Per minute (excl.spaces)	18.330
Words	
Total Typed in Main Document	84
Per minute	2.532
Mean Word Length	6.893
Standard Deviation Word Length	4.731

a) “Process Information” di Matecat

Process Information	
Characters	
Total Typed in Main Document	11604
Total copied	10625
Total typed (incl.spaces)	979
Per minute (incl. spaces)	31.882
Total typed (excl.spaces)	852
Per minute (excl.spaces)	27.746
Words	
Total Typed in Main Document	115
Per minute	3.745
Mean Word Length	7.148
Standard Deviation Word Length	5.983

b) “Process Information” di SmartCAT

Process Information	
Characters	
Total Typed in Main Document	12241
Total copied	10100
Total typed (incl.spaces)	2141
Per minute (incl. spaces)	45.733
Total typed (excl.spaces)	1874
Per minute (excl.spaces)	40.030
Words	
Total Typed in Main Document	252
Per minute	5.383
Mean Word Length	7.270
Standard Deviation Word Length	5.421

c) “Process Information” di SYSTRANet

Figura 3.19 “Process information” di Matecat, SmartCAT e SYSTRANet. Il riepilogo mostra il numero totale di caratteri, spazi inclusi o meno digitati (Total typed) e il numero totale di parole digitate (Total typed in Main Document).

Coerentemente con i risultati di cui sopra, con un totale di 252 parole e 2141 caratteri, SYSTRANet (c) si conferma primo anche in questo parametro.

Tuttavia, il dato che più degli altri fornisce informazioni rilevanti circa la qualità dell’output, correlata all’attività registrata da Inputlog, è di sicuro la suddivisione tra attività su tastiera e attività su mouse svolta nel post-editare il prodotto restituito dai tre programmi web-based. Partendo dal presupposto che, come esposto nel capitolo precedente, l’attività su tastiera corrisponde al processo di post-editing vero e proprio, ovvero il momento in cui il post-editor digita il testo da inserire e cancella le parti errate, mentre l’attività su mouse è imputabile all’esigenza di muoversi lungo il testo per andare ad agire su quei punti che necessitano dell’intervento del post-editor,

una netta demarcazione delle due diverse attività è da vedersi come indice dell'effettiva e quantificabile necessità di intervenire sul testo. In tal senso, i dati esposti in figura 3.20 sembrano confermare suddetta ipotesi e, indirettamente, quanto affermato finora.

Writing Mode	
Keyboard	00:08:54
Total Time (s)	534.460
Number of Clusters Keyboard	165
Mean Cluster Time Keyboard	3.240
Standard Deviation Cluster Time Keyboard	5.640
Number of M-Bursts Keyboard	178
Mean M-Burst Time Keyboard	2.950
Standard Deviation M-Burst Time Keyboard	5.590
Switches Mouse to Keyboard	165
Mouse	00:19:08
Total Time (s)	1.148.670
Number of Clusters Mouse	166
Mean Cluster Time Mouse	6.920
Standard Deviation Cluster Time Mouse	8.950
Number of M-Bursts Mouse	226
Mean M-Burst Time Mouse	3.850
Standard Deviation M-Burst Time Mouse	4.600
Switches Keyboard to Mouse	165

Writing Mode	
Keyboard	00:10:53
Total Time (s)	653.470
Number of Clusters Keyboard	181
Mean Cluster Time Keyboard	3.610
Standard Deviation Cluster Time Keyboard	6.740
Number of M-Bursts Keyboard	203
Mean M-Burst Time Keyboard	2.930
Standard Deviation M-Burst Time Keyboard	5.420
Switches Mouse to Keyboard	181
Mouse	00:14:09
Total Time (s)	849.460
Number of Clusters Mouse	183
Mean Cluster Time Mouse	4.640
Standard Deviation Cluster Time Mouse	5.390
Number of M-Bursts Mouse	222
Mean M-Burst Time Mouse	3.140
Standard Deviation M-Burst Time Mouse	3.120
Switches Keyboard to Mouse	182

a) "Writing mode" di Matecat

b) "Writing Mode" di SmarCAT

Writing Mode	
Keyboard	00:19:53
Total Time (s)	1.193.500
Number of Clusters Keyboard	304
Mean Cluster Time Keyboard	3.930
Standard Deviation Cluster Time Keyboard	4.810
Number of M-Bursts Keyboard	336
Mean M-Burst Time Keyboard	3.280
Standard Deviation M-Burst Time Keyboard	3.830
Switches Mouse to Keyboard	304
Mouse	00:16:32
Total Time (s)	992.090
Number of Clusters Mouse	307
Mean Cluster Time Mouse	3.230
Standard Deviation Cluster Time Mouse	3.340
Number of M-Bursts Mouse	353
Mean M-Burst Time Mouse	2.390
Standard Deviation M-Burst Time Mouse	1.880
Switches Keyboard to Mouse	306

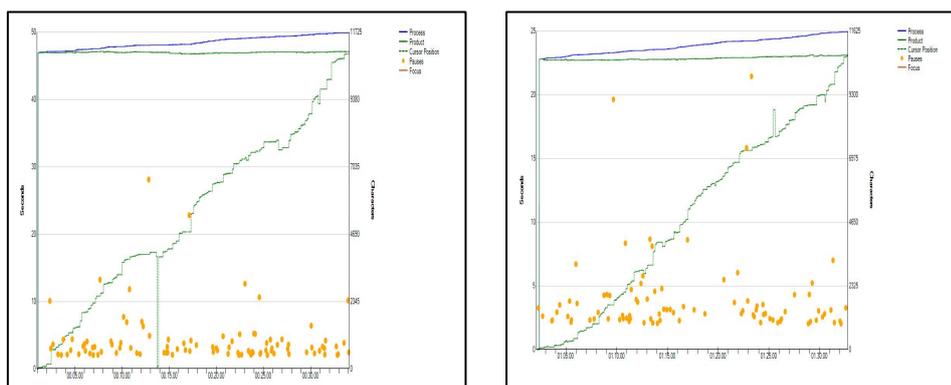
c) "Writing mode" di SYSTRANet

Figura 3.20 Writing Mode di Matecat, SmartCAT e SYSTRANet. Il riepilogo contiene, separatamente, informazioni quali tempo e numero di eventi circa l'attività su tastiera e mouse.

Come è possibile osservare, mentre in Matecat e SmartCAT il tempo dedicato all'attività su mouse eccede in maniera più o meno considerevole quello dedicato all'attività su tastiera, in SYSTRANet si verifica la situazione opposta, alla luce di

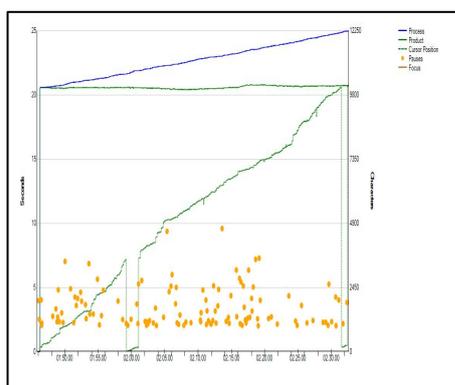
una quantità maggiore di interventi richiesti dal post-editor, inversamente proporzionali, pertanto, alla qualità dell'output del tool.

Infine, si forniscono a seguire i grafici elaborati da Inputlog contenenti la rappresentazione grafica dei dati appena esposti.



a) “Process Graph” di Matecat

b) “Process Graph” di SmartCAT



c) “Process Graph” di SYSTRANet

Figura 3.21 “Process Graph” di Matecat, SmartCAT e SYSTRANet. Il grafico include i seguenti valori: la produzione del testo (linea continua blu), la lunghezza progressiva del documento in corso di scrittura (linea continua verde), movimento del cursore lungo il testo in corso di scrittura (linea tratteggiata verde) e le pause (punto giallo). Per quanto concerne le pause, vengono considerate tali tutte quelle interruzioni che eccedono la soglia stabilita, ovvero 2000 ms, valore di default di Inputlog.

Dall'immagine, il dato che cattura subito l'attenzione del lettore è sicuramente la frequenza con cui si registrano le pause durante il processo di scrittura nei tre testi, prendendo nota dell'addensamento di queste nel grafico relativo al processo di lavorazione sull'output di SYSTRANet. Per quanto concerne la linea tratteggiata

verde, rappresentante il movimento del cursore indicante, su schermo, la posizione in cui verrà inserito il testo digitato, è possibile osservare un andamento costante crescente nel grafico relativo a SmartCAT, mentre negli altri due grafici, soprattutto in SYSTRANet, si osserva un andamento meno lineare, dove repentini picchi decrescenti segnalano il retrocedere del cursore a posizioni precedenti, segnalando quindi una fase di revisione, data dalla necessità dell'individuo di revisionare quanto già scritto.

3.4.2 DYNAMIC QUALITY FRAMEWORK E MATECAT: CALCOLO DELLA PRODUTTIVITÀ E DEL POST-EDITING EFFORT (PEE)

Il Dynamic Quality Framework sviluppato da TAUS, come si è avuto modo di vedere, consente di valutare la qualità di una traduzione, sia essa il prodotto di traduzione automatica o traduzione. Un'altra interessante funzionalità è la misurazione della produttività durante l'attività di post-editing dell'output della machine translation. Come accennato varie volte, DQF consente l'implementazione, tramite API key, di questa funzionalità in Matecat. I dati restituiti da DQF consentono di avere un'idea della qualità della traduzione automatica in funzione della produttività misurata, in termini di numero di parole post-editate per ora, tempo totale impiegato e tempo medio per segmento e così via.

Alcuni dei dati elaborati e presentati da DQF sono condivisi anche da Inputlog, motivo per cui sarà utile confrontarli al fine di individuare e valutare eventuali discrepanze. Tuttavia, un dato di grande utilità, non esaminabile in Inputlog, è la distanza che intercorre tra la MT e la sua versione post-editata. DQF, tramite l'algoritmo di Levenshtein, consente di ottenere e valutare questo dato. In figura 3.22 è possibile osservare i risultati restituiti circa questo dato da DQF.

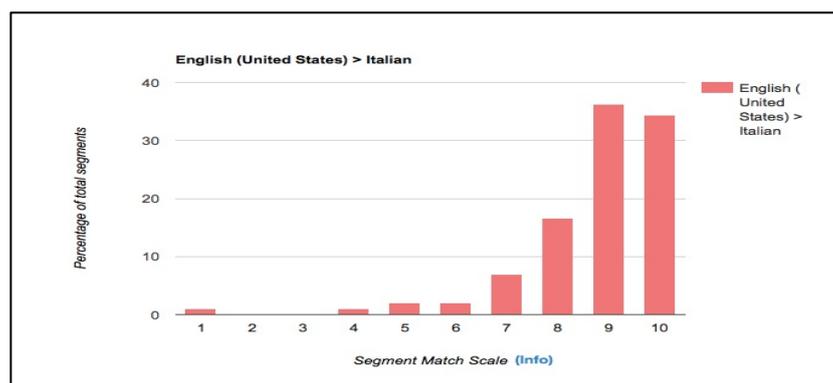


Figura 3.22 Edit Distance Graph dell'output di Matecat. L'asse x mostra il livello di match tra l'output della machine translation e l'output post-editato. I valori tra 1 e 5 indicano una corrispondenza bassa o quasi inesistente, necessitando un gran quantitativo di correzione o addirittura una traduzione *from scratch*, mentre i valori superiori a 5 indicano match più alti, dove 10 corrisponde a un full match. L'asse y mostra la percentuale di output della MT compreso entro quei valori.

Come suggerisce il grafico, la percentuale di output della MT richiedente massicci interventi di correzione o ritraduzione è piuttosto bassa, mentre la stragrande maggioranza ricade tra i valori 9 e 10. Di conseguenza, salvo alcuni segmenti che richiedevano interventi consistenti, sussumendo questi dati è possibile affermare che l'output di Matecat si è rivelato essere di buona qualità, comportando un vantaggio produttivo non indifferente, sebbene sia opportuno ricordare che, data la tipologia testuale, la percentuale di segmenti identici, autopropagatisi dopo l'intervento sul primo segmento della serie, era piuttosto consistente, un dato che ha certamente influenzato positivamente la percentuale di segmenti richiedenti un'intensa attività di post-editing.

Un altro dato interessante emerso dall'analisi della produttività condotta da DQF sulla prestazione di Matecat è il calcolo della produttività media, estratto dal calcolo del numero medio di parole post-editate per ora sulla base della prestazione analizzata da DQF. Questo dato (WPH), coerente tra l'altro con le stime generali su cui Matecat basa la propria analisi al momento della creazione del progetto di traduzione, può risultare di sicura utilità nel calcolo del tempo necessario a completare un ipotetico progetto di traduzione, qualora Matecat venisse implementato in un contesto professionale. Di seguito si fornisce il grafico elaborato da DQF relativo al numero medio di parole post-editate per ora.

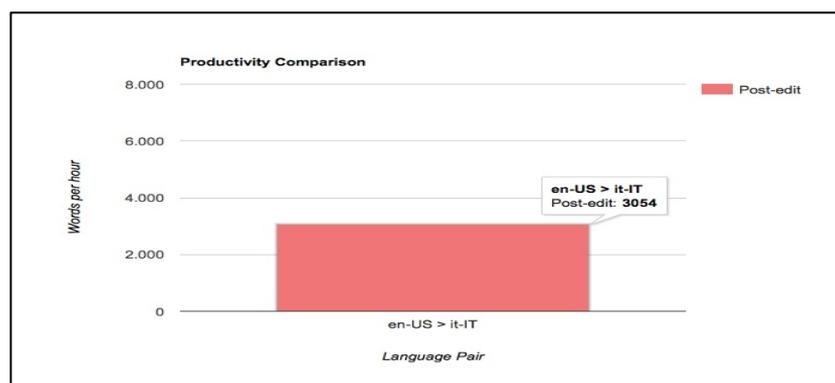


Figura 3.23 Produttività media registrata in Matecat.

Infine, come detto in precedenza, Matecat stesso dispone di un editing log che registra l'attività condotta su Matecat e calcola il tempo totale necessario all'esecuzione e al completamento dell'attività, il numero di secondi medio impiegato per parola e lo sforzo medio (*Post-Editing Effort*) prodotto sommariamente nell'arco di tempo totale trascorso dall'inizio alla fine del progetto. A seguire si forniscono i dati elaborati dall'editing log di Matecat.

The screenshot shows the 'Editing Log' for project 'test_1'. It includes a notice about DQF being used for statistics and a table of statistics for translation job 259495 - en-US > it-IT. A 'Get Support' button is visible on the right side.

Words	Total Time-to-edit	Your avg secs/word	Your avg PEE
1583	00h:36m:47s	You're fast! 1.4s (avg 11.25s)	12.65% (avg 52.16%)

Figura 3.24 Editing log di Matecat.

Il dato più interessante, poiché consente una correlazione con quelli estratti da Inputlog, è il tempo totale impiegato per completare il progetto. Come si può osservare, questo dato evidenzia uno scarto di circa 3 minuti in più nella prestazione di Matecat, dato inputabile a un'interfaccia utente talvolta carente in termini di chiarezza e funzionalità, nonché un sistema di gestione e implementazione dei tag che, in più di un'occasione, risulta dispersivo e non funzionale all'attività svolta dall'utente, senza contare inoltre che, come si puntualizzerà nelle discussioni in chiusura al capitolo, una volta esportata la traduzione ultimata, diversi considerevoli problemi di layout sono da imputarsi proprio a tali tag.

3.5 DISCUSSIONE DEI RISULTATI

I risultati restituiti dalla metodologia di valutazione applicata sono stati utili a definire, grazie a dati pratici ottenibili solo tramite prove empiriche ed esperienza diretta, la qualità dell'output dei tre sistemi presi in esame in questa sede. Il primo dato che emerge dall'analisi di suddetti risultati è il fatto che, seppure tramite metodologie diverse, tutti i risultati sembrano aver puntato nella stessa direzione. Infatti, sebbene stilare una classifica di rendimento dell'output dei tre tool sia un'operazione riduttiva e non in grado di cogliere le sostanziali differenze nelle prestazioni dei sistemi MT tra un parametro e l'altro, è tuttavia possibile concludere che, sommariamente, SmartCAT si è rivelato essere il tool più performante dei tre, mentre SYSTRANet è risultato essere quello meno efficiente, ponendo quindi nel mezzo tra i due la prestazione globale di Matecat. È evidente come una classifica di questo genere non renda giustizia, ad esempio, della maggiore cura, rispetto a Matecat, posta da SYSTRANet per quanto concerne le regole grammaticali e sintattiche. Ciononostante, al fine di individuare il tool la cui implementazione nel flusso di lavoro di un traduttore può portare i vantaggi maggiori, è opportuno individuare, in maniera piuttosto diretta, qual è quello che ha necessitato di minore tempo per il complemento del task assegnato, quello avente il maggior numero di segmenti il cui output poteva essere effettivamente riutilizzato per essere riveduto e corretto, elevandone la qualità al livello publishable, e infine quello avente uno stile il più vicino possibile a quello richiesto dalla tipologia testuale adottata. Ciò premesso, è possibile concludere che questo tool è SmartCAT.

Prima di arrivare a una conclusione, è tuttavia doveroso spendere alcune parole circa alcune considerazioni che non hanno trovato spazio nei paragrafi precedenti. Una fra queste riguarda l'effetto che la tipologia testuale adottata ha avuto sulle prestazioni dei tre tool. Come accennato nel paragrafo 3.2, il testo legislativo comunitario è stato scelto in questa sede coerentemente alle scelte testuali fatte in studi ed esperimenti simili a questo e di cui abbonda la letteratura in materia, al fine di testare le prestazioni di diversi tool su questa tipologia testuale tanto oggetto di analisi, piuttosto che quello del singolo. È opportuno ricordare al lettore, tuttavia, che il lavoro di coredazione ad opera degli organi preposti delle versioni parallele dei documenti prodotti in seno all'UE difficilmente trova eguali, motivo per cui la

reperibilità di testi comunitari già tradotti è quanto mai sconfinata, tanto il plurilinguismo è stato la base della fondazione dell'Unione europea. In virtù di questa condizione, i tool presi in esame si trovano a disporre di un asset che non ha virtualmente limiti e che può incidere in maniera nettamente positiva sulle prestazioni di questi.

D'altro canto, è altrettanto doveroso puntualizzare che le problematiche riscontrate nella gestione dei formati originali dei documenti ha inficiato sulle prestazioni dei tre tool. Sebbene Matecat e SmartCAT siano in grado di elaborare file PDF, come detto in precedenza, gli effetti sulla segmentazione del documento si sono rivelati spesso problematici, condizionando pesantemente il rendimento dei tool.

Una considerazione a parte merita anche il salvataggio delle traduzioni nel formato nativo una volta ultimate e revisionate e, più nello specifico, del layout finale del documento tradotto. Si è già parlato profusamente delle difficoltà inerenti alla gestione di differenti formati riscontrate dai tre tool. Al di là delle diverse difficoltà traduttive che possono venirsi a creare dalla segmentazione del documento, un ulteriore problema nasce quando layout e caratteri del documento non risultano essere adattati a quello del documento source. Il tool che ha generato il maggior numero di problemi circa questo aspetto è stato Matecat, per il quale nel documento tradotto salvato come file di testo nativo è stato riscontrato l'alternarsi frequente di font e grandezza del carattere diversi. Un tipo di problema del genere, in un contesto professionale, richiederebbe un intervento massiccio di revisione non compatibile con le tempistiche imposte nel mercato da traduzione automatica e post-editing, interventi non eseguibili tra l'altro internamente al tool ma richiedente un editor di file di testo quale può essere Word. Nell'output degli altri due tool, dal canto loro, non si segnalano problemi di questa natura.

Infine, sebbene questa situazione sia già stata esposta precedentemente, è opportuno sottolineare nuovamente l'incompatibilità di fondo tra questi tool e programmi di revisione esterni a questi, quale Xbench. Per quanto, come suggerito, i problemi incontrati nell'utilizzazione di Xbench fossero risolvibili, questi avrebbero comunque richiesto procedimenti intermedi e dispendiosi in termini di tempo ed energie, senza contare la necessità di dover ricorrere a ulteriori programmi di

localizzazione quali Studio, il cui utilizzo può non essere garantito a tutti, soprattutto nei contesti lavorativi in cui questi tool open-source mirano a imporsi. Per tale motivo, si sente il bisogno di auspicare una maggiore futura compatibilità con suddetti programmi esterni.

Conclusione

Nell'introduzione a questo studio si è fatto notare come le innovazioni in campo tecnologico abbiano condizionato le attuali modalità di lavoro in diversi settori professionali. In questa sede ci si è concentrati sull'ormai indissolubile rapporto che la tecnologia ha intessuto nel corso degli anni con la professione del traduttore. Se, anni fa, l'idea di servirsi di programmi informatici e di strumenti quali memorie di traduzione e glossari terminologici fosse percepito come qualcosa di rivoluzionario, attualmente il supporto che un computer può fornire al traduttore è ancora maggiore. Questa opportunità, con il tempo, è stata accolta, compresa e sfruttata dall'industria, la quale ha visto nella traduzione automatica la possibilità di incrementare produttività e profitti, tenendo così il passo con le nuove esigenze dettate dall'odierna maggiore fruibilità di contenuti, multimediali e non, accessibili in rete. Questa tendenza è facilmente individuabile nel mercato, nel quale si assiste allo sviluppo, pubblicizzazione e messa a disposizione di servizi gratuiti o comunque a costi ridotti, messi a disposizione in formato web-based e open-source.

Alla luce di questa premessa, lo scopo di questa tesi era quello di vedere in che modo, da un punto di vista qualitativo, questi tool fossero un riflesso di questa nuova tendenza e apertura a nuove forme di traduzione e se, pertanto, in questi si può vedere una possibile implementazione e utilizzazione futura in contesti lavorativi professionali. È possibile affermare che l'obiettivo è stato in parte raggiunto. La metodologia di analisi e valutazione adottata ha permesso di mettere in evidenza punti di forza, nonché debolezze, dei tre tool presi a campione. È doveroso affermare che nessuna delle prestazioni dei tool è risultata tuttavia priva di problematiche e ostacoli alla corretta esecuzione di un eventuale progetto di traduzione.

Come esposto in precedenza, l'intrinseca difficoltà di questi tool ad essere associati all'utilizzo di programmi di revisione e QA esterni ha posto problematicità di grossa portata. Alla luce di un comparto di revisione e Quality Assurance non del tutto impeccabile, quando non del tutto assente, la possibilità di fare ricorso a programmi di revisione esterni sembra quasi imprescindibile ma cionondimeno un passaggio difficoltoso e dispendioso in termini di tempo. La strategia adottata in questa sede per aggirare questo ostacolo è stata sì fruttuosa, nonché applicabile in contesti lavorativi veri e propri, ma richiedente un quantitativo di tempo del tutto

incompatibile con gli standard della machine translation e del post-editing. Rinunciare alla revisione condotta esternamente ai suddetti tool sembrerebbe pertanto l'unica opzione percorribile, una variabile che, tuttavia, farebbe da discriminante nella scelta di avvalersi o meno di questa tecnologia, poiché le ripercussioni sulla qualità del prodotto finito potrebbero essere ingenti.

Quello della qualità traduttiva è stato un altro discorso affrontato profusamente. Come detto in precedenza, approntare una metodologia di valutazione consona alla tipologia testuale, allo *skopos* della traduzione e alla tipologia di utenza finale può non essere sempre semplice come obiettivo, poiché, in funzione delle molteplici variabili, i risultati possono pendere tanto da un lato quanto dall'altro e restituirne quindi di falsati. In questa sede pertanto si è approntata una metodologia funzionale allo scopo dello studio: valutare singolarmente l'output *raw* della machine translation dei tre tool al fine di individuare quello più performante, prima come prodotto a sé stante e poi a confronto con una *reference translation*. I diversi parametri di valutazione hanno puntato, tutti in maniera concorde e uniforme, sulle prestazioni superiori di SmartCAT e su quelle meno efficienti di SYSTRANet. Il confronto con una traduzione effettuata senza l'ausilio di CAT tool ha invece confermato la generale tendenza a una maggiore accuratezza, sempre e comunque entro i canoni della MT, nel testo tradotto per quanto concerne alcuni punti del testo giuridico comunitario esaminato, a discapito della resa stilistica, non sempre all'altezza della *reference translation*. Suddetti risultati sono stati confermati dall'analisi su standard produttivi condotta tramite il keystroke tracker Inputlog e l'utility di valutazione della produttività di DQF.

In conclusione, quanto esposto in questo studio dimostra le effettive potenzialità di questa tecnologia e apre la strada a una maggiore implementazione di questi sistemi, auspicando tuttavia maggiori ricerche, approfondimenti e migliorie al fine di garantire prestazioni più solide e affidabili. L'analisi delle performance dei tre tool sottintende un considerevole interesse circa il futuro di questo approccio alla traduzione e l'effettiva possibilità che questi tool attraggano una porzione sempre maggiore di utenti.

Premesso che questo studio ha solamente gettato le basi per un'analisi di questa nuova tendenza, è certo che un'attenzione costante dovrà essere posta su tecnologie

come Matecat e SmartCAT, al fine di analizzare gli sviluppi futuri di questi e la loro risposta in base al feedback della futura utenza. D'altro canto, l'impegno profuso da TAUS nella creazione dei suoi parametri di valutazione in DQF sottintende un importante interesse per l'industria circa la qualità, motivo per cui è ragionevole aspettarsi future ripercussioni circa tale questione in funzione dell'introduzione su mercato di questa tecnologia gratuita e web-based. Dal momento che non è stato possibile avvalersene in questa sede, inoltre, in futuro sarebbe senz'altro interessante e rilevante includere gli eventuali risultati restituibili da un'analisi di movimenti e fissazioni oculari, per confermarne l'effettiva utilità ai fini di questa ricerca, andando a vedere quindi se questi confermano o meno i risultati ottenuti.

Bibliografia

- Aikawa, T., Schwartz, L., King, R., Corston-Oliver, M., & Lozano, C. (2007). Impact of controlled language on translation quality and post-editing in a statistical machine translation environment. *Proceedings of the MT Summit XI, 10-14 September 2007, Copenhagen, Denmark*. Allschwil: EAMT, 1-7.
- Allen, J., & Hogan, C. (2000). Toward the development of a post editing module for raw machine translation output: A controlled language perspective. *Third International Controlled Language Applications Workshop (CLAW-00), 29-30 April 2000, Seattle, WA*. Stroudsburg, PA: Association for Computational Linguistics, 62-71.
- Arnold, D. (1994). *Machine translation: an introductory guide*. Cambridge, MA: Blackwell Pub.
- Banerjee, S., Alon L. (2005). METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. *Proceedings of the ACL workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization, 29 June 2005, Ann Arbor, MI*. Stroudsburg, PA: Association for Computational Linguistics, 65-72.
- Berghoefer, K. (2013). *TAUS Quality Evaluation Summit: Adequacy, Fluency and/or Readability Evaluation*. [video]. Amsterdam: TAUS. Consultato da: <https://www.youtube.com/watch?v=GEVnSPUkPxY>
- Brown, P. F., Cocke, J., Pietra, S. A. D., Pietra, V. J. D., Jelinek, F., Lafferty, J. D. & Roossin, P. S. (1990). "A statistical approach to machine translation". *Computational linguistics*, 16(2): 79-85. Consultato da: <http://www-cgi.cs.cmu.edu/afs/cs.cmu.edu/user/lafferty/www/pub/mtcl90.ps>
- Callison Burch, C. (2007). *Decoding in machine statistical translation*. Baltimore, MD: John Hopkins University.
- Carl, M., Kay, M. & Jensen, K. T. H. (2010). Long Distance Revisions in Drafting and Post-editing. *Special issue: Natural Language Processing and its Applications*, 46: 193-2014. Consultato da: <http://pics.cicling.org/2010/Vol46.pdf#page=203>
- Cosmai, D. (2007). *Tradurre per l'Unione europea. Prassi, problemi e prospettive del multilinguismo comunitario dopo l'ampliamento a est*. Milano: Hoepli.

- De Almeida, G., & O'Brien, S. (2010). Analysing post-editing performance: correlations with years of translation experience. *Proceedings of the 14th annual conference of the European association for machine translation, 27-28 May 2010, St. Raphaël, France*. Allschwil: EAMT, 27-28.
- Doddington, G. (2002). Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. *Proceedings of the second international conference on Human Language Technology Research, 24-27 March 2002, San Diego, CA*. San Francisco CA: Morgan Kaufmann Publishers Inc., 138-145.
- Doherty, S., O'Brien, S. & Carl, M. (2010) "Eye tracking as an MT evaluation technique". *Machine translation*, 24(1): 1-13. Consultato da: http://openarchive.cbs.dk/bitstream/handle/10398/8045/SubmissionforMT_dohertyobriencarl.pdf
- Dorr, B. J. (1987). UNITRAN: an interlingual approach to machine translation. *Proceedings of the sixth National conference on Artificial intelligence, 13-17 July 1987, Seattle, WA*. Palo Alto, CA: AAAI Press, 534-539.
- Dorr, B. J., (1990). A cross-linguistic approach to translation. *Proceedings of the Third International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Language, 11-13 June 1990, Austin, TX*. Austin, TX: Linguistic Research Centre, 13-33.
- Dorr, B. J., (1992). A parameterized approach to integrating aspect with lexical-semantics for machine translation. *Proceedings of the 30th annual meeting on Association for Computational Linguistics, 28 June - 2 July 1992, Newark, NJ*. Stroudsburg, PA: Association for Computational Linguistics, 257-264.
- Dorr, B. J. (1993). *Machine translation: a view from the Lexicon*. Cambridge, MA: MIT press.
- Fiederer, R. & O'Brien, S. (2009). "Quality and machine translation: a realistic objective". *The Journal of Specialised Translation*, 11: 52-74. Consultato da: http://www.jostrans.org/issue11/art_fiederer_obrien.php
- Germann, U., Jahr, M., Knight, K., Marcu, D. & Yamada, K. (2001). Fast decoding and optimal decoding for machine translation. *Proceedings of the 39th Annual Meeting on Association for Computational Linguistics, 6-11 July 2001, Toulouse, France*. Stroudsburg, PA: Association for Computational Linguistics, 228-235.

- González-Rubio, J. (2015). "On the Effective Deployment of Current Machine Translation Technology". *Localisation Focus*, 14(2): 28-41. Consultato da: <https://www.localisation.ie/locfocus/issues/14/2#>
- Gough, N., & Way, A. (2003). Controlled generation in example-based machine translation. *Proceedings of MT Summit IX, 23-28 September 2003, New Orleans, LA*. Dublin: DORAS.
- Gough, N., & Way, A. (2004). Example-based controlled translation. *Proceedings of EAMT-04 - 9th Workshop of the European Association for Machine Translation, 26-27 April 2004, Valetta, Malta*. Dublin: DORAS.
- Green, S., Heer, J. & Manning, C. D. (2013). The efficacy of human post-editing for language translation. *Proceedings of the SIGCHI conference on human factors in computing systems, 27 April - 2 May 2013, Paris, France*. New York, NY: ACM, 439-448.
- Hutchins, W. J. (2001). "Machine translation over fifty years". *Histoire épistémologie langage*, 23(1): 7-31. Consultato da: http://www.persee.fr/doc/hel_0750-8069_2001_num_23_1_2815
- Hutchins, W. J. (2006) "Machine Translation: History". In B. Keith. *Encyclopedia of Language & Linguistics* (2^a ed.) (Vol 7). Oxford: Elsevier, 375-383.
- Hutchins, W. J. (2007). "Machine translation: A concise history". *Journal of Translation Studies*, 13: 29-70. Consultato da: <http://www.hutchinsweb.me.uk/CUHK-2006.pdf>
- Hutchins, W. J. (2014). "Machine translation: A brief history". In E.F.K. Koerner & R.E. Asher. *Concise history of the language sciences: from the Sumerians to the cognitivists*. Oxford, Elsevier, , 431-445.
- Knight, K. & Chander, I. (1994). Automated postediting of documents. *Proceedings of AAAI - 94, 31 July - 4 August, Seattle, WA*. Palo Alto, CA: AAAI, 779-784.
- Lopez, A. (2008). *Statistical machine translation*. New York, NY: ACM Computing Surveys (CSUR).
- Marcu, D. (2015). "At the boundary between humans and machines", *Fair Trade Translation*. Consultato da: <https://www.linkedin.com/pulse/boundary-between-humans-machines-daniel-marcu>

- Marcu, D. (2015). "Translation Trends: a 2015 TAUS Conference Perspective", *Fair Trade Translation*. Consultato da: <https://www.linkedin.com/pulse/translation-trends-2015-taus-conference-perspective-daniel-marcu>
- Mori, L. (2002). "Variazioni interlinguistiche, creazione terminologica e dinamismo comunicativo nei documenti comunitari". *Romaneske, Associazione dei Romanisti dell'Università Cattolica di Leuven*, 4(27): 56-74.
- Mori, L. (2008). "Gli anni Sessanta e la costruzione dell'identità linguistica europea. Sulla formazione della varietà comunitaria d'italiano". In M. De Pasquale, G. Dotoli, M. Selvaggio. *I linguaggi del Sessantotto. Atti del Convegno multidisciplinare*. Roma: Editrice Apes, 531-544.
- Muegge, U. (2008). *An introduction to the use of controlled language*. Muegge.cc Consultato da: <http://www.muegge.cc/controlled-language.htm>
- O'Brien, S. (2002). Teaching post-editing: a proposal for course content. *Proceedings of the 6th EAMT Workshop Teaching Machine Translation, 14-15 November 2002, Manchester, United Kingdom*. Allschwil: EAMT, 99-106.
- O'Brien, S. (2003). Controlling controlled english. An analysis of several controlled language rule sets. *Proceedings of EAMT-CLAW 2003, 15-17 May 2003, Dublin, Ireland*. Allschwil: EAMT, 105-114.
- O'Brien, S. (2006). "Pauses as indicators of cognitive effort in post-editing machine translation output". *Across Languages and Cultures*, 7(1): 1-21. DOI: 10.1556/ACROSS.7.2006.1.
- O'Brien, S. (2011). "Towards predicting post-editing productivity". *Machine translation*, 25(3): 197-215. DOI: 10.1007/s10590-011-9096-7.
- O'Brien, S. (2013). *TAUS Quality Evaluation Summit: Error Typology Quality Evaluation*. [video]. Amsterdam: TAUS. Consultato da: <https://www.youtube.com/watch?v=C1kiR7YfQrM>
- Och, F. J. (2003). Minimum error rate training in statistical machine translation. *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics, 7-12 July 2003, Sapporo, Japan*. Stroudsburg, PA: Association for Computational Linguistics, 160-167.
- Offersgaard, L., Povlsen, C., Almsten, L. K., & Maegaard, B. (2008). Domain specific MT in use. *Proceedings of EAMT 2008: 12th annual conference of the*

- European Association for Machine Translation, 22-23 September 2008, Hamburg, Germany.* Allschwil: EAMT, 150-159.
- Papineni, K., Roukos, S., Ward, T. & Zhu, W. J. (2002). BLEU: a method for automatic evaluation of machine translation. *Proceedings of the 40th annual meeting on association for computational linguistics, 7-12 July 2002, Philadelphia, PA.* Stroudsburg PA: Association for Computational Linguistics, 311-318.
- Rascu, E. (2006). A controlled language approach to text optimization in technical documentation. *Proceedings of KONVENS, 4-7 October 2006, Konstanz, Germany.* Köln: DGfS, 107-114.
- Ryan, Joann P. (1988). "The Role of the Translator in Making an MT System Work: Perspective of a Developer". In M. Vasconcellos. *Technology as translation strategy.* Philadelphia, PA: John Benjamins Publishing, 127-132.
- Sabatini, F. (1999). "Rigidità-esplicitzza" vs "elasticità-implicitzza": possibili parametri massimi per una tipologia dei testi. In F., Sabatini, G., Skytte. *Linguistica Testuale Comparativa*, (vol. 42), Copenhagen: Museum Tusulanum Forlag, 141-172.
- Scarpa, F. (2008). *La traduzione specializzata.* Milano: Hoepli.
- Schwenk, H., Barrault L., Frédéric B., Bougares F., Hazem A., & Servan, C. (2015). MateCat-An Open Source CAT Tool with closely integrated User Specific Statistical Machine Translation. *Proceedings of the 1st TAO-CAT, 18-20 June 2015, Angers, France.*
- Séguinot, C. L. C. (1989). *The translation process.* Toronto: HG Publications.
- Snover, M., Dorr, B., Schwartz, R., Micciulla, L. & Makhoul, J. (2006). A study of translation edit rate with targeted human annotation. *Proceedings of association for machine translation in the Americas, 8-12 August 2006, Cambridge, MA.* Stroudsburg, PA: AMTA, 223-231.
- Specia, L., Saunders, C., Turchi, M., Wang, Z., & Shawe-Taylor, J. (2009). Improving the confidence of machine translation quality estimates. *Proceedings of the Twelfth Machine Translation Summit (MT Summit XII), 26-30 August 2009, Ottawa, Canada.* Stroudsburg, PA: AMTA, 136-143.
- Specia, L., Turchi, M., Cancedda, N., Dymetman, M., & Cristianini, N. (2009). Estimating the sentence-level quality of machine translation systems. *Proceedings*

- of the 13th Conference of the European Association for Machine Translation, 14-15 May 2009, Barcelona, Spain. Allschwil: EAMT, 28-37.*
- Specia, L., & Farzindar, A. (2010). Estimating machine translation post-editing effort with HTER. *Proceedings of the Second Joint EM+/CNGL Workshop Bringing MT to the User: Research on Integrating MT in the Translation Industry (JEC 10), 4th November 2010, Denver, CO. Stroudsburg, PA: AMTA, 33-41.*
- TAUS (2012). *Controlled language.* Consultato da: https://www.taus.net/knowledgebase/index.php?title=Controlled_Natural_Languages
- TAUS (2013). *Translation Technology Showcase Webinar: Gengo & Translated.* [webinar]. Amsterdam: TAUS. Consultato da: <https://events.taus.net/events/webinars/webinar-october-2-2013-gengo-translated>
- TAUS (2015). *Translation Technology Showcase Webinar: Coach & SmartCAT.* [webinar]. Amsterdam: TAUS. Consultato da: <https://events.taus.net/events/webinars/translation-technology-webinar-may-6-2015-coach-smartcat#smartcat>
- TAUS (2016). *Translation Technology Showcase Webinar: Tauyou.* [webinar]. Amsterdam: TAUS. Consultato da: <https://www.taus.net/academy/videos#taus-translation-technology-showcase-webinars>
- TAUS (n.d.). *Confidence estimation.* Consultato da: https://www.taus.net/knowledgebase/index.php?title=Confidence_estimation (data di consultazione 10/01/2016).
- TAUS (n.d.). *Automatic evaluation.* Consultato da: https://www.taus.net/knowledgebase/index.php?title=Automatic_evaluation (data di consultazione 20/01/2016).
- TAUS (n.d.). *Error typology.* Consultato da: https://www.taus.net/knowledgebase/index.php?title=Error_typology (data di consultazione: 20/01/2016).
- TAUS (n.d.). *Rules of Post-editing.* Consultato da: https://www.taus.net/knowledgebase/index.php?title=Rules_of_Post-editing (data di consultazione 26/01/2016).

- Unione Europea (2011). *Manuale interistituzionale di convenzioni redazionali*. Bruxelles: Ufficio delle pubblicazioni.
- Van Slype, G. (1979). Critical study of methods for evaluating the quality of machine translation. Prepared for the Commission of European Communities Directorate General Scientific and Technical Information and Information Management. Bruxelles, Bureau Marcel van Dijk. Report BR 19142.
- Vasconcellos, M. (1988). *Technology as translation strategy*. Philadelphia, PA: John Benjamins Publishing.
- Wang, Y.Y. & Waibel A. (1997). Decoding algorithm in statistical machine translation. *Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics, 7-12 July 1997, Madrid, Spain*. Stroudsburg, PA: Association for Computational Linguistics, 366-372.

Appendici

A seguire, si allegano nelle appendici il testo originale della proposta congiunta e la *reference translation*, nonché gli output dei tre tool analizzati. Per non alterarne il formato e consentirne la visione così come sono nella loro versione originale, si è scelto di non adattarli allo stile del resto della tesi, motivo per cui presenteranno impaginazione e layout differente.

Appendice A: Joint proposal for a Council decision (JOIN/2015/35/FINAL)



**HIGH REPRESENTATIVE
OF THE UNION FOR
FOREIGN AFFAIRS AND
SECURITY POLICY**

Brussels, 17.12.2015
JOIN(2015) 35 final

2015/0303 (NLE)

Joint Proposal for a

COUNCIL DECISION

**on the signing, on behalf of the European Union, of the Cooperation Agreement on
Partnership and Development between the European Union and the Islamic Republic of
Afghanistan**

EXPLANATORY MEMORANDUM

1. CONTEXT OF THE PROPOSAL

The Decision proposed constitutes the legal instrument for authorising the signature of the Cooperation Agreement on Partnership and Development between the European Union and the Islamic Republic of Afghanistan ('the Agreement').

In July 2011, the Foreign Affairs Council expressed '*its willingness to negotiate an ambitious and balanced EU-Afghanistan agreement reflecting its long-term commitment to the country's development as well as the principles and conditions on which the future partnership will be based*'. It invited the EEAS and the Commission to draft a negotiation mandate for a cooperation agreement ahead of the Bonn Conference,¹ held on 5 December 2011. In November 2011, the Council adopted a Decision authorising the Commission and the High Representative to negotiate a Cooperation Agreement on Partnership and Development (CAPD) with Afghanistan². Three rounds of negotiations were held over the year that followed (the last ending in November 2012). After a break of over two years, talks were resumed in 2015, with the new Afghan Government. The fourth and final round of negotiations took place in Brussels on 29 April 2015 and was concluded successfully. The two sides initialled the Agreement in Kabul on 2 July 2015 in the presence of President Ghani.

The Agreement is the first contractual relationship between the European Union and Afghanistan and underpins the EU's commitment to supporting Afghanistan's future development during its 'decade of transformation' agreed at the Bonn conference in 2011. By strengthening political dialogue and improving cooperation in a broad range of areas, the Agreement consolidates the European Union's engagement with Afghanistan. It acknowledges the results of the international conferences on Afghanistan held in Bonn, Chicago, Kabul, Tokyo and London.

The Agreement includes provisions on political dialogue and on cooperation in a broad range of areas. It draws on the EU's standard political clauses on human rights and the International Criminal Court, and includes commitments related to the rights of women and children. The Agreement builds on the principles of mutual accountability and reiterates the willingness of the parties to address shared concerns, including: 1) the fight against terrorism, international crime and illegal trafficking; 2) non-proliferation, disarmament and nuclear security; 3) Weapons of Mass Destructions (WMD); 4) Small Arms and Light Weapons (SALW); and 5) counter-narcotics. The provisions on cooperation cover the following sectors: infrastructure development, energy, transport, health, natural resources, tax, education and culture, employment and social affairs, science and technology, and environment and climate change. The Agreement also emphasises the importance of legal cooperation and affirms the parties' commitment to fighting organised crime, money laundering and corruption.

2. RESULTS OF CONSULTATIONS WITH INTERESTED PARTIES AND IMPACT ASSESSMENTS

The Council has been kept informed at all stages of the negotiations. It has been regularly consulted in the relevant Working Party on Asia.

The European Parliament has also been kept fully informed throughout negotiations and the High Representative of the Union for Foreign Affairs and Security Policy (High

¹ Foreign Affairs Council Conclusions of 18 July 2011 (doc. st12865/11)

² Council Decisions of 10 November 2011 (doc. st 16146/11 and st16147/11)

Representative') sent a copy of the initialled agreement to the President of the European Parliament on.... ..

The High Representative and the Commission consider that the objectives set by the Council in its negotiating directives have been met, and that the draft Agreement can be submitted for signature.

3. LEGAL ELEMENTS OF THE PROPOSAL

Legal basis

According to settled case-law, the choice of the legal basis must rest on objective factors amenable to judicial review, which include the aim and content of that measure.

The aim of the Agreement, as described in Article 2 *Nature and Scope*, is to establish a partnership between the parties that rests on a multipillar approach, in order to strengthen dialogue and cooperation.

The Agreement covers political cooperation (Title II), development cooperation (Title III), cooperation on trade and investment (Title IV), cooperation in matters of freedom, security and justice (Title V), sectoral cooperation (Title VI) and regional cooperation (Title VII). In addition, the Agreement includes provisions on the institutional set-up (Title VIII) and final provisions (Title IX).

A detailed analysis of the aim and content of the Agreement indicates that some of the provisions of the Agreement fall within the scope of the common foreign and security policy, and some other elements fall under the trade policy, as well as under the development cooperation policy of the European Union. The analysis further shows that neither of these elements can be regarded as ancillary to any of the other mentioned elements and none of the three mentioned elements can be clearly identified as the main component. Consequently, the proposal should be based on multiple legal bases, namely Article 37 TEU and Articles 207 and 209 TFEU.

Legal nature

The legal bases referred to above confer power on the EU to enter into agreements with third parties in the areas that the Agreement covers. In particular:

- Article 37 TEU grants the EU the power to conclude agreements with third parties in the area of common foreign and security policy.
- Article 207 TFEU grants the EU the power to conclude agreements in the area of common commercial policy. Article 3(1) TFEU gives the EU exclusive competence in this area.
- Article 209 TFEU grants the EU the power to conclude agreements with non-EU countries in the area of development cooperation policy.

Pursuant to case-law, the question of whether a particular provision from an agreement *comes within the competence of the [Union] is one which relates to the attribution and, thus, the very existence of that competence, and not to its exclusive or shared nature*'. Therefore, and to the extent that the Treaties have attributed competence to the Union in a policy area, the Union may exercise such competence externally.

Pursuant to Article 3(2) TFEU, the EU also has exclusive competence for the conclusion of an international agreement *'in so far as its conclusion may affect common rules or alter their scope'*. The Court of Justice has held in a ruling that negotiations *may affect common rules or alter their scope* to the extent that *the content of the negotiations [...] 'falls within an area covered to a large extent by common EU rules'... Therefore those negotiations fall within the exclusive competence of the European Union"*. Recent case-law also clarifies that any analysis of a competence, especially an exclusive competence, must take into account not

only the existing Union rules but also the *'foreseeable future development [...] of those rules'*.

The Treaties have thus conferred power on the EU to act in all the areas covered by the Agreement.

Other legal considerations

The Agreement sets up an institutional framework composed of the Joint Committee (see Title VIII, Article 49 *Institutional framework*). The Joint Committee may set up special committees or working groups to assist it with the performance of its tasks. Article 54 *Fulfilment of obligations* also sets out a procedure for settling disputes, to be used should one of the parties fail to fulfil its obligations under the Agreement.

The Agreement is valid for an initial period of ten years. It will automatically be extended for subsequent consecutive periods of five years unless, six months prior to the expiry of its validity, one of the parties notifies the other in writing of its intention not to extend it. The Agreement can be terminated at six months' notice.

Joint Proposal for a

COUNCIL DECISION

on the signing, on behalf of the European Union, of the Cooperation Agreement on Partnership and Development between the European Union and the Islamic Republic of Afghanistan

THE COUNCIL OF THE EUROPEAN UNION,

Having regard to the Treaty on European Union, in particular Article 37 thereof,

Having regard to the Treaty on the Functioning of the European Union, in particular Articles 207 and 209, in conjunction with Article 218(5) and the second subparagraph of Article 218(8) thereof,

Having regard to the proposal from the European Commission and the High Representative of the Union for Foreign Affairs and Security Policy,

Whereas:

- (1) In November 2011, the Council authorised the Commission and the High Representative of the Union for Foreign Affairs and Security Policy to open negotiations with the Islamic Republic of Afghanistan on a Cooperation Agreement on Partnership and Development³.
- (2) Negotiations on the Cooperation Agreement were successful, and the Agreement was initialled on 2 July 2015 in Kabul.
- (3) The Agreement should therefore be signed on behalf of the EU, subject to its conclusion at a later date,

HAS ADOPTED THIS DECISION:

Article 1

The signing of the Cooperation Agreement on Partnership and Development between the European Union and the Islamic Republic of Afghanistan is hereby approved on behalf of the Union, subject to the conclusion of the said Agreement.

The text of the Agreement is attached to this Decision.

Article 2

The Council General Secretariat shall establish the instrument of full powers to sign the Agreement, subject to its conclusion, for the person(s) indicated by the negotiators of the Agreement.

Article 3

This Decision shall enter into force on the day following that of its adoption.

³ Council Decisions of 10 November 2011 (doc. st 16146/11 and st16147/11)

Done at Brussels,

*For the Council
The President*

Appendice B: reference translation



ALTO RAPPRESENTANTE
DELL'UNIONE PER
GLI AFFARI ESTERI E
LA POLITICA DI SICUREZZA

Bruxelles, 17.12.2015
JOIN(2015) 35 final

2015/0303 (NLE)

Proposta congiunta di

DECISIONE DEL CONSIGLIO

**relativa alla firma, a nome dell'Unione europea, dell'accordo di cooperazione sul
partenariato e lo sviluppo tra l'Unione europea e la Repubblica islamica di Afghanistan**

RELAZIONE

1. CONTESTO DELLA PROPOSTA

La decisione proposta costituisce lo strumento giuridico che autorizza la firma dell'accordo di cooperazione sul partenariato e lo sviluppo tra l'Unione europea e la Repubblica islamica di Afghanistan ("l'accordo").

Nel luglio 2011, il Consiglio «Affari esteri» ha espresso *“la sua volontà a negoziare un ambizioso ed equilibrato accordo UE-Afghanistan che rifletta il suo impegno a lungo termine nello sviluppo del paese, nonché i principi e le condizioni su cui si baserà il futuro partenariato”*. Esso ha invitato il SEAE e la Commissione a redigere un mandato negoziale per un accordo di cooperazione in vista della conferenza di Bonn,¹ tenutasi il 5 dicembre 2011. Nel novembre 2011, il consiglio ha adottato una decisione che autorizza la Commissione e l'alto rappresentante a negoziare un accordo di cooperazione sul partenariato e lo sviluppo con l'Afghanistan.² Si sono tenuti tre cicli di negoziati nel corso dell'anno seguente (dei quali l'ultimo si è concluso a novembre 2012). Dopo un'interruzione di più di due anni, il dialogo è ripreso nel 2015, con il nuovo governo afgano. Il quarto e ultimo ciclo di negoziati ha avuto luogo a Bruxelles il 29 aprile 2015 e si è concluso con successo. L'accordo è stato siglato dalle due parti a Kabul il 2 luglio 2015 in presenza del presidente Ghani.

L'accordo costituisce il primo rapporto contrattuale tra l'Unione europea e l'Afghanistan e promuove l'impegno dell'UE nel supportare il futuro sviluppo dell'Afghanistan durante il suo “decennio di trasformazione” concordato nella conferenza di Bonn nel 2011. Rafforzando il dialogo politico e migliorando la cooperazione in un'ampia gamma di settori, l'accordo consolida l'impegno dell'Unione europea con l'Afghanistan. Lo stesso riconosce i risultati delle conferenze internazionali che hanno avuto luogo a Bonn, Chicago, Kabul, Tokyo e Londra.

L'accordo include disposizioni sul dialogo politico e sulla cooperazione in un'ampia gamma di settori. Esso si ispira alle clausole politiche standard dell'UE in materia di diritti umani e della Corte penale internazionale, e include impegni relativi ai diritti dei minori e delle donne. L'accordo si basa sui principi di responsabilità reciproca e ribadisce la volontà delle parti ad affrontare preoccupazioni comuni, tra cui: 1) la lotta contro il terrorismo, la delinquenza internazionale e il traffico illecito; 2) la non proliferazione, il disarmo e la sicurezza nucleare; 3) armi di distruzione di massa (ADM); 4) armi leggere e di piccolo calibro; 5) la lotta alla droga. Le disposizioni in materia di cooperazione comprendono i seguenti settori: sviluppo delle infrastrutture, energia, trasporti, sanità, risorse naturali, fiscalità, istruzione e cultura, occupazioni e affari sociali, scienza e tecnologia, e ambiente e cambiamenti climatici. L'accordo evidenzia altresì l'importanza della cooperazione giuridica e ribadisce l'impegno delle parti alla lotta contro la criminalità organizzata, il riciclaggio di denaro e la corruzione.

2. CONSULTAZIONE DELLE PARTI INTERESSATE E VALUTAZIONI D'IMPATTO

Il Consiglio è stato informato in tutte le fasi dei negoziati. È stato regolarmente consultato nel gruppo di lavoro competente sull'Asia.

Il Parlamento europeo è stato altresì pienamente informato durante le negoziazioni e l'alto rappresentante dell'Unione per gli affari esteri e la politica di sicurezza (“alto

¹ Conclusioni del Consiglio «Affari esteri» del 18 luglio 2011 (doc. st12865/11)

² Decisione del Consiglio del 10 novembre 2011 (doc. st 16146/11 e st16147/11)

rappresentante”) ha inviato una copia dell’accordo siglato al presidente del Parlamento europeo su... ..

L’alto rappresentante e la Commissione ritengono che gli obiettivi fissati dal Consiglio nelle sue direttive di negoziato siano stati realizzati e che il progetto di accordo possa essere presentato alla firma.

3. ELEMENTI GIURIDICI DELLA PROPOSTA

Base giuridica

Secondo una giurisprudenza costante, la scelta della base giuridica deve basarsi su elementi oggettivi suscettibili di sindacato giurisdizionale, tra cui figurano lo scopo e il contenuto dell’atto.

L’accordo, descritto nell’articolo 2 *Natura e portata*, ha lo scopo di istituire un partenariato tra le parti risiede in un approccio multipilastro per rafforzare il dialogo e la cooperazione.

L’accordo comprende la cooperazione politica (Titolo II), cooperazione allo sviluppo (Titolo III), cooperazione in commercio e investimenti (Titolo IV), cooperazione in materia di libertà, sicurezza e giustizia (Titolo V), cooperazione settoriale (Titolo VI) e cooperazione regionale (Titolo VII). Inoltre, l’accordo prevede disposizioni sugli assetti istituzionali (Titolo VIII) e disposizioni finali (Titolo IX).

Un’analisi dettagliata dello scopo e del contenuto dell’accordo indica che alcune delle disposizioni dell’accordo sono contemplate dalla politica estera e di sicurezza comune, e altri elementi rientrano nella politica commerciale, nonché nella politica di cooperazione allo sviluppo dell’Unione europea. L’analisi altresì rivela che nessuno dei suddetti elementi può essere considerato accessorio agli altri elementi menzionati e nessuno dei tre elementi menzionati può essere chiaramente identificato come l’elemento principale. Di conseguenza, è opportuno che la proposta sia basata su molteplici basi giuridiche, ossia l’articolo 37 del TUE e gli articoli 207 e 209 del TFUE.

Natura giuridica

Le basi giuridiche succitate conferiscono all’UE il potere di concludere accordi con paesi terzi nei settori contemplati dall’accordo. In particolare:

- l’articolo 37 del TUE conferisce all’UE il potere di concludere accordi con paesi terzi nel settore della politica estera e di sicurezza comune;
- l’articolo 207 del TFUE conferisce all’UE il potere di concludere accordi in materia di politica commerciale comune. A norma dell’articolo 3, paragrafo 1, del TFUE l’Unione europea ha competenza esclusiva in questo settore;
- l’articolo 209 del TFUE conferisce all’UE il potere di concludere accordi con paesi terzi nel settore della politica di cooperazione allo sviluppo.

Secondo la giurisprudenza, stabilire se una disposizione di un accordo *rientri nella competenza dell’[Unione]* è una questione che riguarda l’attribuzione e, dunque, l’esistenza stessa della competenza, e non la sua natura esclusiva o ripartita”. Per tale motivo, e nella misura in cui i trattati hanno riconosciuto la competenza dell’Unione in un determinato settore, l’Unione può esercitare tale competenza esternamente.

A norma dell’articolo 3, paragrafo 2, del TFUE, l’UE ha inoltre competenza esclusiva per la conclusione di accordi internazionali *"nella misura in cui tale conclusione può incidere su norme comuni o modificarne la portata."* In una sentenza la Corte di giustizia ha statuito che i negoziati *possono incidere sulle norme comuni dell’Unione o modificarne la portata* nella misura in cui *il contenuto dei negoziati [...] “rientra in un settore ampiamente disciplinato*

da norme comuni dell'Unione"... Pertanto, detti negoziati rientrano nella competenza esclusiva dell'Unione". La giurisprudenza recente chiarisce inoltre che un'analisi relativa alla competenza, in particolare alla competenza esclusiva, deve tener conto non solo delle norme dell'Unione in vigore, ma anche delle "prevedibili prospettive di evoluzione [...] di tali norme".

I trattati hanno pertanto conferito all'Unione europea il potere di agire in tutti i settori disciplinati dall'accordo.

Altre considerazioni giuridiche

L'accordo stabilisce un quadro istituzionale composto dal Comitato misto (cfr. Titolo VIII, articolo 49 *Quadro istituzionale*). Il Comitato misto può istituire comitati speciali o gruppi di lavoro che siano di ausilio allo svolgimento dei propri compiti. L'articolo 54 *Adempimento degli obblighi* dispone inoltre una procedura per la risoluzione delle controversie, da utilizzare qualora una delle parti non adempia i propri obblighi a norma dell'accordo.

L'accordo è valido per un periodo iniziale di dieci anni. Esso è automaticamente rinnovato per successivi consecutivi periodi di cinque anni a meno che, sei mesi prima della cessazione della sua validità, una delle parti notifichi per iscritto all'altra l'intenzione di non rinnovarlo. L'accordo può essere denunciato con un preavviso di sei mesi.

Proposta congiunta di

DECISIONE DEL CONSIGLIO

relativa alla firma, a nome dell'Unione europea, dell'accordo di cooperazione sul partenariato e lo sviluppo tra l'Unione europea e la Repubblica islamica di Afghanistan

IL CONSIGLIO DELL'UNIONE EUROPEA:

visto il trattato sull'Unione europea, in particolare l'articolo 37,

visto il trattato sul funzionamento dell'Unione europea, in particolare gli articoli 207 e 209, in combinato disposto con l'articolo 218, paragrafo 5, e l'articolo 218, paragrafo 8, secondo comma,

vista la proposta della Commissione europea e dell'Alto rappresentante dell'Unione per gli affari esteri e la politica di sicurezza,

considerando quanto segue:

- (1) Nel novembre 2011, il Consiglio ha autorizzato la Commissione e l'alto rappresentante dell'Unione per gli affari esteri e la politica di sicurezza ad avviare i negoziati con la Repubblica islamica di Afghanistan per un accordo di cooperazione sul partenariato e lo sviluppo³.
- (2) I negoziati per l'accordo di cooperazione sono andati a buon fine e l'accordo è stato siglato il 2 luglio 2015 a Kabul.
- (3) È dunque opportuno che l'accordo sia firmato a nome dell'Unione, in attesa della sua conclusione in una data successiva,

HA ADDOTATO LA SEGUENTE DECISIONE:

Articolo 1

È approvata la firma, a nome dell'Unione europea, dell'accordo di cooperazione sul partenariato e lo sviluppo tra l'Unione europea e la Repubblica islamica di Afghanistan, con riserva della conclusione di detto accordo.

Il testo dell'accordo da firmare è accluso alla presente decisione.

Articolo 2

Il segretariato generale del Consiglio definisce lo strumento dei pieni poteri per la firma dell'accordo, con riserva della sua conclusione, per la persona o le persone indicate dai negoziatori del medesimo.

Articolo 3

³ Decisione del Consiglio del 10 novembre 2011 (doc. st 16146/11 e st16147/11)

La presente decisione entra in vigore il giorno successivo all'adozione.
Fatto a Bruxelles, il

*Per il Consiglio,
il presidente*

Appendice C: output di Matecat



ALTO RAPPRESENTANTE
nell'Unione
AFFARI ESTERI E
POLITICA DI SICUREZZA

Bruxelles, 2015/12/17
JOIN (2015) 35 definitivo
2015/0303 (NLE)

Proposta congiunta di

Decisione del Consiglio

**relativa alla firma, a nome dell'Unione europea, dell'accordo di
cooperazione sul partenariato e lo sviluppo tra l'Unione
europea e la Repubblica islamica dell'Afghanistan**

RELAZIONE

1. CONTESTO DELLA PROPOSTA

La proposta di decisione costituisce lo strumento giuridico per autorizzare la firma dell'accordo di cooperazione sul partenariato e lo sviluppo tra l'Unione europea e la Repubblica islamica dell'Afghanistan ('l'accordo').

Nel luglio 2011, il Consiglio Affari esteri ha espresso *'la sua volontà di negoziare un ambizioso ed equilibrato accordo -Afghanistan UE che riflette il suo impegno a lungo termine per lo sviluppo del paese, così come i principi e le condizioni su cui si baserà il futuro partenariato'*. Esso invitato il SEAE e la Commissione a elaborare un mandato negoziale per un accordo di cooperazione in vista della Conferenza di Bonn,¹ del 5 dicembre 2011. In Novembre 2011 il Consiglio ha adottato una decisione che autorizza la Commissione e l'Alto rappresentante per negoziare un accordo di cooperazione sul partenariato e lo sviluppo (DPCA) con l'Afghanistan². Tre cicli di negoziati si sono svolti nel corso dell'anno che ha seguito (l'ultima finale a novembre 2012). Dopo una pausa di oltre due anni, i colloqui sono ripresi nel 2015, con il nuovo governo afgano. Il quarto e ultimo ciclo di negoziati ha avuto luogo a Bruxelles il 29 aprile 2015 e si è conclusa con successo. Le due parti hanno siglato l'accordo a Kabul il 2 luglio 2015 alla presenza del Presidente Ghani.

L'accordo è il primo rapporto contrattuale tra l'Unione europea e l'Afghanistan e rafforza l'impegno dell'UE a sostenere lo sviluppo futuro dell'Afghanistan durante la sua 'decennio di trasformazione' concordato alla conferenza di Bonn nel 2011. Con il rafforzamento del dialogo politico e migliorare la cooperazione in un'ampia gamma di settori, l'accordo consolida l'impegno dell'Unione europea con l'Afghanistan. Riconosce i risultati delle conferenze internazionali sull'Afghanistan tenutasi a Bonn, Chicago, Kabul, Tokyo e Londra.

L'accordo contiene disposizioni in materia di dialogo politico e di cooperazione in una vasta gamma di settori. Essa si basa su clausole politiche standard dell'UE in materia di diritti umani e la Corte penale internazionale, e comprende gli impegni relativi ai diritti delle donne e dei bambini. L'accordo si basa sui principi della responsabilità reciproca e ribadisce la volontà delle parti di affrontare problemi comuni, tra cui: 1) la lotta contro il terrorismo, la criminalità internazionale e il traffico illegale; 2) la non proliferazione, il disarmo e la sicurezza nucleare; 3) Weapons of Mass Distruzioni (WMD); 4) di piccolo calibro e armi leggere (SALW); e 5) la lotta al narcotraffico. Le disposizioni in materia di cooperazione riguardano i seguenti settori: lo sviluppo delle infrastrutture, l'energia, i trasporti, la salute, le risorse naturali, tasse, istruzione e cultura, occupazione e affari sociali, scienza e tecnologia, e l'ambiente e il cambiamento climatico. L'accordo sottolinea inoltre l'importanza della cooperazione giuridica e afferma l'impegno delle parti a combattere la criminalità organizzata, il riciclaggio di denaro e la corruzione.

2. RISULTATI DELLE CONSULTAZIONI CON LE PARTI INTERESSATE E VALUTAZIONI DELL'IMPATTO

Il Consiglio è stato tenuto informato in tutte le fasi dei negoziati. E 'stato regolarmente consultato nel relativo gruppo di lavoro per l'Asia.

Il Parlamento europeo è stato anche tenuto pienamente informato per tutta negoziati e la Alto rappresentante dell'Unione per gli affari esteri e la politica di sicurezza ('High

¹ Consiglio Affari Esteri conclusioni del 18 luglio 2011 (doc. St12865 / 11)

² decisioni del Consiglio del 10 novembre 2011 (doc. 16146/11 st e st16147 / 11)

Rappresentante ') ha inviato una copia dell'accordo siglato al Presidente del Parlamento europeo e

Parlamento europeo in data

L'alto rappresentante e la Commissione ritengono che gli obiettivi fissati dal Consiglio nelle sue direttive di negoziato siano stati raggiunti e che il progetto di accordo possono essere presentati per la firma.

3. ELEMENTI GIURIDICI DELLA PROPOSTA

Base giuridica

Secondo la giurisprudenza costante, la scelta della base giuridica deve basarsi su elementi oggettivi, suscettibili di sindacato giurisdizionale, tra cui lo scopo e il contenuto di tale misura.

L'obiettivo dell'accordo, come descritto all'articolo 2, *Natura e campo di applicazione*, È quello di stabilire una relazione tra le parti che si basa su un approccio multipilastro, al fine di rafforzare il dialogo e la cooperazione.

L'accordo riguarda la cooperazione politica (titolo II), la cooperazione allo sviluppo (titolo III), cooperazione in materia di commercio e investimenti (titolo IV), cooperazione in materia di libertà, sicurezza e giustizia (titolo V), cooperazione settoriale (titolo VI) e la cooperazione regionale (titolo VII). Inoltre, l'accordo comprende disposizioni relative al assetto istituzionale (titolo VIII) e disposizioni finali (Titolo IX).

Un'analisi dettagliata dello scopo e del contenuto dell'accordo indica che alcune disposizioni della caduta accordo nell'ambito della politica estera e di sicurezza comune, e alcuni altri elementi rientrano nella politica commerciale, così come nel quadro della politica di cooperazione allo sviluppo dell'Unione europea. L'analisi mostra inoltre che nessuno di questi elementi possono essere considerate accessorie a ciascuno degli altri elementi menzionati e nessuno dei tre elementi menzionati possono essere chiaramente identificate come componente principale. Di conseguenza, la proposta dovrebbe essere basata su più basi giuridiche, vale a dire l'articolo 37 TUE e articoli 207 e 209 del TFUE.

Natura giuridica

Le basi giuridiche di cui al potere conferire sopra l'Unione europea a stipulare accordi con terzi nei settori che l'accordo copre. In particolare:

- Articolo 37 del TUE concede all'UE il potere di concludere accordi con terzi in materia di politica estera e di sicurezza comune.
- L'articolo 207 del TFUE conferisce all'UE il potere di concludere accordi in materia di politica commerciale comune. Articolo 3 (1) TFUE dà competenza esclusiva dell'UE in questo settore.
- L'articolo 209 del TFUE conferisce all'Unione europea il potere di concludere accordi con paesi terzi nel settore della politica di cooperazione allo sviluppo.

Ai sensi della giurisprudenza, la questione se una disposizione particolare da un accordo *arriva sono di competenza della [dell'Unione] è una questione che riguarda l'attribuzione e, dunque, l'esistenza stessa di tale competenza, e non la sua natura esclusiva o ripartita*. Pertanto, e nella misura in cui i trattati hanno attribuito competenza all'Unione in un settore politico, l'Unione può esercitare tale competenza esternamente.

Ai sensi dell'articolo 3 (2), del TFUE, l'Unione europea dispone di una competenza esclusiva per la conclusione di un accordo internazionale *'nella misura in cui la sua conclusione possa*

incidere su norme comuni o alterarne la scope '. La Corte di giustizia ha affermato in una sentenza che i negoziati *può incidere su norme comuni o alterarne la portata* nella misura in cui *il contenuto dei negoziati [...] 'in uno dei settori coperta in gran parte da norme comuni dell'UE'... Pertanto tali negoziati è di competenza esclusiva dell'Unione europea* ". Giurisprudenza recente chiarisce, inoltre, che qualsiasi analisi di una competenza, in particolare una competenza esclusiva, deve tener conto non solo delle norme dell'Unione esistenti, ma anche i *'lo sviluppo futuro prossimo [...] di quelle regole* ".

I trattati hanno così conferito potere all'Unione europea di agire in tutti i settori contemplati dall'accordo.

Altre considerazioni giuridiche

L'accordo istituisce un quadro istituzionale composto dal comitato misto (vedi titolo VIII, l'articolo 49 *Quadro istituzionale*). La commissione mista può istituire commissioni speciali o gruppi di lavoro per assisterlo con le prestazioni dei suoi compiti. Articolo 54 *Adempimento degli obblighi* stabilisce anche una procedura di risoluzione delle controversie, per essere usato se una delle parti non abbia rispettato gli obblighi previsti dall'accordo.

L'accordo è valido per un periodo iniziale di dieci anni. Esso sarà automaticamente prorogata per successivi periodi consecutivi di cinque anni, salvo, sei mesi prima della scadenza della sua validità, una delle parti notifichi all'altra per iscritto la sua intenzione di non estenderla. L'accordo può essere denunciato in preavviso di sei mesi.

Proposta congiunta di

Decisione del Consiglio

relativa alla firma, a nome dell'Unione europea, dell'accordo di cooperazione sul partenariato e lo sviluppo tra l'Unione europea e la Repubblica islamica dell'Afghanistan

IL CONSIGLIO DELL'UNIONE EUROPEA,

visto il trattato sull'Unione europea, in particolare l'articolo 29,

visto il trattato sul funzionamento dell'Unione europea, in particolare l'articolo 207 e l'articolo 352, in combinato disposto con l'articolo 218, paragrafo 5, e paragrafo 8, secondo comma,

vista una proposta congiunta della Commissione europea e dell'Alto rappresentante dell'Unione per gli affari esteri e la politica di sicurezza,

considerando quanto segue:

- (1) Nel novembre 2011, il Consiglio ha autorizzato la Commissione e l'Alto rappresentante dell'Unione per gli affari esteri e la politica di sicurezza ad avviare negoziati con la Repubblica islamica di Afghanistan su un accordo di cooperazione sul partenariato e lo sviluppo.³
- (2) I negoziati per l'accordo di cooperazione hanno avuto successo, e l'accordo è stato siglato il 2 luglio 2015 a Kabul.
- (3) Occorre che l'accordo sia firmato a nome dell'Unione, con riserva della sua conclusione in una data successiva,

HA ADOTTATO LA PRESENTE DECISIONE:

Articolo 1

La firma dell'accordo di cooperazione sul partenariato e lo sviluppo tra l'Unione europea e la Repubblica islamica dell'Afghanistan è approvato a nome dell'Unione, con riserva della conclusione di tale accordo.

Il testo dell'accordo è accluso alla presente decisione.

Articolo 2

Il segretariato generale del Consiglio definisce lo strumento dei pieni poteri a firmare l'accordo, con riserva della sua conclusione, per la persona (s) indicato dai negoziatori dell'accordo.

³ Le decisioni del Consiglio del 10 novembre 2011 (doc. 16146/11 st e st16147 / 11)

Articolo 3

La presente decisione ha effetto il giorno successivo alla data dell'adozione.

Fatto a Bruxelles, addì

Per il Consiglio

Il Presidente

Appendice D: output di SmartCAT



EUROPEO
COMMISSION

ALTO RAPPRESENTANTE
LA POLITICA DI
GUASTI DI AFFARI DELL'UNIONE
SPESEDIGESTIONE PER LE TRANSAZIONI

Bruxelles,
17.12.2015
JOIN(2015) 35 def.

Congiunta proposta per un

DECISIONE DEL CONSIGLIO

**relativa alla firma, a nome dell'Unione europea, della
onPartnership di accordo di cooperazione e sviluppo tra l'Unione europea e la Repubblica
islamica di
Afghanistan**

1. CONTESTO DELLA PROPOSTA

La decisione proposta costituisce lo strumento giuridico che autorizza la firma dell'accordo di cooperazione sul partenariato e lo sviluppo tra l'Unione europea e la Repubblica islamica dell'Afghanistan («l'accordo»).

Nel luglio 2011, il Consiglio affari esteri ha espresso *«la volontà di negoziare un accordo UE-Afghanistan ambizioso ed equilibrato che riflettono l'impegno a lungo termine lo sviluppo del paese, nonché i principi e le condizioni su cui si baserà il futuro partenariato. Invita il SEAE e la Commissione a elaborare un mandato negoziale per un accordo di cooperazione in vista della conferenza di Bonn,¹ tenutasi il 5 dicembre 2011. Nel novembre 2011, il Consiglio ha adottato una decisione che autorizza la Commissione e l'alto rappresentante per negoziare un accordo di cooperazione sul partenariato e lo sviluppo (CAPD) con l'Afghanistan.² Tre round di negoziati si sono svolte nel corso dell'anno che seguì (l'ultimo conclusosi nel novembre 2012). Dopo una pausa di oltre due anni, i colloqui sono stati ripresi nel 2015, con il nuovo governo afgano. Il quarto e ultimo round di negoziati ha avuto luogo a Bruxelles il 29 aprile 2015 e si è concluso con successo. I due lati siglato l'accordo a Kabul il 2 luglio 2015 alla presenza del Presidente Ghani.*

L'accordo è il primo rapporto contrattuale tra l'Unione europea e l'Afghanistan ed è alla base dell'UE di impegno a sostenere lo sviluppo futuro dell'Afghanistan durante il suo 'decennio di trasformazione' approvato alla conferenza di Bonn nel 2011. Rafforzamento del dialogo politico e migliorando la cooperazione in una vasta gamma di aree, l'accordo consolida impegno dell'Unione europea in Afghanistan. Essa riconosce i risultati delle conferenze internazionali sull'Afghanistan, tenute a Bonn, Chicago, Kabul, Tokyo e Londra.

L'accordo contiene disposizioni sul dialogo politico e sulla cooperazione in una vasta gamma di settori. Disegna su clausole politiche standard dell'UE sui diritti umani e la Corte penale internazionale e comprende impegni relazionati ai diritti di donne e bambini. L'accordo si basa sui principi della responsabilità reciproca e ribadisce la volontà delle parti di affrontare preoccupazioni comuni, tra cui: 1) la lotta contro il terrorismo, la criminalità internazionale e il traffico illegale; 2) non-proliferazione, disarmo e sicurezza nucleare; 3) le armi di distruzione di massa (WMD); 4) armi di piccolo calibro e leggere (SALW); e 5) contro il narcotraffico. Le disposizioni sulla cooperazione coprono i seguenti settori: sviluppo delle infrastrutture, energia, trasporti, salute, risorse naturali, fiscale, educazione e cultura, occupazione e affari sociali, scienza e tecnologia e ambiente e cambiamento climatico. L'accordo, inoltre, sottolinea l'importanza della cooperazione giuridica e afferma l'impegno delle parti a combattere la criminalità organizzata, il riciclaggio di denaro e la corruzione.

2. RISULTATI DELLE CONSULTAZIONI CON LE PARTI INTERESSATE E VALUTAZIONE D'IMPATTO

Il Consiglio è stato informato in tutte le fasi dei negoziati. Esso è stato consultato regolarmente in gruppo di lavoro pertinenti sull'Asia.

Il Parlamento europeo ha anche stato pienamente informato nel corso dei negoziati e l'alto rappresentante dell'Unione per gli affari esteri e la politica di sicurezza (' alta

¹ Conclusioni del Consiglio affari esteri del 18 luglio 2011 (doc. st12865/11)

² decisioni del Consiglio del 10 novembre 2011 (doc. st 16146/11 e st16147/11)

Rappresentante ') una copia dell'accordo parafato ha inviato al Presidente del Parlamento europeo

L'alto rappresentante e la Commissione ritengono che siano stati soddisfatti gli obiettivi fissati dal Consiglio nelle sue direttive di negoziato, e che il progetto di accordo possa essere presentato per la firma.

3. Base giuridica elementi giuridici della proposta

Secondo una costante giurisprudenza, la scelta della base giuridica deve basarsi su elementi oggettivi, suscettibili di sindacato giurisdizionale, tra i quali figurano lo scopo e il contenuto di tale misura.

L'obiettivo dell'accordo, come descritto in articolo 2 *natura e la portata*, è di stabilire una relazione tra le parti che si basa su un approccio multipillar, al fine di rafforzare il dialogo e la cooperazione.

L'accordo copre la cooperazione politica (titolo II), la cooperazione allo sviluppo (titolo III), la cooperazione in materia di commercio e investimenti (titolo IV), cooperazione in materia di libertà, sicurezza e giustizia (titolo V), cooperazione settoriale (titolo VI) e cooperazione regionale (titolo VII). Inoltre, l'accordo include disposizioni sull'assetto istituzionale (titolo VIII) e finali (titolo IX).

Un'analisi dettagliata dell'obiettivo e il contenuto dell'accordo indica che alcune delle disposizioni della caduta accordo nell'ambito della politica estera e politica di sicurezza e alcuni altri elementi rientrano nell'ambito della politica commerciale, così come sotto la politica di cooperazione allo sviluppo dell'Unione europea. L'analisi ulteriore dimostra che nessuno di questi elementi può essere considerata accessoria ad uno qualsiasi degli altri menzionati elementi e nessuno dei tre elementi menzionati possono essere chiaramente identificati come componente principale. Di conseguenza, la proposta dovrebbe essere basata su più basi giuridiche, vale a dire articolo 37 del tue e articoli 207 e 209 del TFUE.

Natura giuridica

Le basi giuridiche di cui sopra conferiscono potere all'UE di stipulare accordi con terzi nelle aree che l'accordo copre. In particolare:

3. Articolo 37 borse TEU all'UE il potere di concludere accordi con terzi nella zona di politica estera e di sicurezza comune.
4. Articolo 207 sovvenzioni del TFUE l'Unione europea il potere di concludere accordi nel settore della politica commerciale comune. Articolo 3, paragrafo 1 del TFUE conferisce all'UE la competenza esclusiva in questo settore.
5. Articolo 209 sovvenzioni del TFUE l'Unione europea il potere di concludere accordi con paesi extracomunitari nell'area della politica di cooperazione allo sviluppo.

Ai sensi della giurisprudenza, la domanda di se una disposizione particolare da un accordo rientra la competenza del [Unione] è uno che riguarda l'attribuzione e, quindi, l'esistenza stessa di tale competenza e non alla sua natura esclusiva o condivisa '. Di conseguenza, e nella misura in cui i trattati hanno attribuito competenza all'Unione in un settore politico, l'Unione può esercitare tale competenza esternamente.

Ai sensi dell'articolo 3, paragrafo 2 del TFUE, l'UE ha inoltre competenza esclusiva per la conclusione di un accordo internazionale *'nella misura in cui può incidere su norme comuni o alterarne la portata'*. La Corte di giustizia ha tenuto in una sentenza che negoziati *possono incidere su norme comuni o alterarne la portata* nella misura in cui *il contenuto dei negoziati [...]' cade all'interno di un'area*

coperto in larga misura da norme comuni... Pertanto tali negoziati rientrano nella competenza esclusiva dell'Unione europea". Giurisprudenza recente chiarisce, inoltre, che qualsiasi analisi di una competenza, soprattutto una competenza esclusiva, devono tener conto non solo le regole dell'Unione esistente, ma anche la *'prevedibile sviluppo futuro [...] di tali norme ''*.

I trattati hanno conferito così potere all'UE di agire in tutti i settori contemplati dall'accordo.

Altre considerazioni legali

Il set di accordi su un quadro istituzionale composta del Comitato misto (Vedi titolo VIII, articolo 49 *istituzionale*). Il Comitato misto può creare comitati speciali o gruppi di lavoro per assisterlo con le prestazioni dei suoi compiti. Articolo che 54 *adempimento degli obblighi* stabilisce inoltre una procedura per la risoluzione delle controversie, da utilizzarsi se una delle parti non riescono per inadempimento ai sensi dell'accordo.

L'accordo è valido per un periodo iniziale di dieci anni. Verrà automaticamente prorogato per successivi periodi di cinque anni consecutivi a meno che, sei mesi prima della scadenza della sua validità, una delle parti notificherà a altra per iscritto della sua intenzione di non estenderla. L'accordo può essere denunciato in preavviso di sei mesi.

Congiunta proposta per un

DECISIONE DEL CONSIGLIO

**relativa alla firma, a nome dell'Unione europea, della
onPartnership di accordo di cooperazione e sviluppo tra l'Unione europea e la Repubblica
islamica di
Afghanistan**

IL CONSIGLIO DELL'UNIONE EUROPEA,

Visto il trattato sull'Unione europea, in particolare l'articolo 37,

Visto il trattato sul funzionamento dell'Unione europea, in particolare gli articoli 207 e 209, in combinato disposto con l'articolo 218 5 e il secondo comma dell'articolo 218(8)

Vista la proposta della Commissione europea e l'alto rappresentante dell'Unione per gli affari esteri e la politica di sicurezza,

Considerando che:

- (1) Nel novembre 2011, il Consiglio ha autorizzato la Commissione e l'alto rappresentante dell'Unione per gli affari esteri e la politica di sicurezza per aprire i negoziati con la Repubblica islamica dell'Afghanistan, su un accordo di cooperazione sul partenariato e lo sviluppo.³
- (2) I negoziati sull'accordo di cooperazione hanno avuto successo, e l'accordo è stato siglato il 2 luglio 2015 a Kabul.
- (3) L'accordo dovrebbe pertanto essere firmato a nome dell'UE, fatta salva la sua conclusione in una data successiva,

HA ADOTTATO LA PRESENTE DECISIONE:

Articolo 1

La firma dell'accordo di cooperazione sul partenariato e lo sviluppo tra l'Unione europea e la Repubblica islamica dell'Afghanistan è approvata a nome dell'Unione, fatta salva la conclusione di detto accordo.

Il testo dell'accordo è accluso alla presente decisione.

Articolo 2

Segretariato generale del Consiglio istituisce lo strumento di pieni poteri a firmare l'accordo, fatta salva la sua conclusione, per la o le persone indicate dai negoziatori dell'accordo.

³ Decisioni del Consiglio del 10 novembre 2011 (doc. st 16146/11 e st16147/11)

Articolo 3

La presente decisione entra in vigore il giorno successivo a quello della sua adozione. Fatto a Bruxelles,

*Per il Consiglio
il Presidente*

RELAZIONE

1. CONTESTO DELLA PROPOSTA

La decisione proposta costituisce lo strumento giuridico per l'autorizzazione della firma dell'accordo di cooperazione sull'associazione e dello sviluppo fra l'Unione Europea e la Repubblica islamica di Afghanistan ('l'accordo').

Nel luglio 2011, il Consiglio di affari esteri ha espresso il *'la sua volontà di negoziare un accordo ambizioso ed equilibrato di UE-Afghanistan che riflette il suo impegno a lungo termine allo sviluppo del paese come pure i principi e le circostanze su cui l'associazione futura sarà basata'*. Ha invitato il EEAS e la Commissione a disegnare un mandato di negoziato per un accordo di cooperazione davanti alla conferenza di Bonn,¹ tenuto il 5 dicembre 2011. Nel novembre 2011, il Consiglio ha adottato una decisione che autorizza la Commissione e l'alto rappresentante a negoziare un accordo di cooperazione sull'associazione e sullo sviluppo (CAPD) con Afghanistan². Tre giri dei negoziati sono stati mantenuti l'anno che è seguito (l'ultimo termine nel novembre 2012). Dopo che una rottura in due anni, colloqui è stata ripresa nel 2015, con il nuovo governo afgano. Il quarto e giro dei negoziati finale ha avuto luogo a Bruxelles il 29 aprile 2015 ed è stato concluso con successo. I due lati hanno siglato l'accordo in Kabul il 2 luglio 2015 in presenza di presidente Ghani.

L'accordo è la prima relazione contrattuale fra l'Unione Europea e Afghanistan e sostiene l'impegno dell'UE a sostenere lo sviluppo futuro di Afghanistan durante la sua decade del *'di trasformazione'* acconsentita alla conferenza di Bonn nel 2011. Rinforzando il dialogo politico e migliorando la cooperazione in una vasta gamma di aree, l'accordo consolida l'impegno dell'Unione Europea con Afghanistan. Riconosce i risultati delle conferenze internazionali su Afghanistan ha tenuto a Bonn, Chicago, Kabul, Tokyo e Londra.

L'accordo comprende le disposizioni sul dialogo politico e sulla cooperazione in una vasta gamma di aree. Attinge le clausole politiche standard dell'UE sui diritti dell'uomo e sulla Corte penale internazionale e comprende gli impegni relativi ai diritti delle donne e dei bambini. Le configurazioni di accordo per principi di responsabilità reciproca e ripete la volontà dei partiti alle preoccupazioni condivise indirizzato, includente: 1) la lotta contro il terrorismo, il crimine internazionale ed il traffico illegale; 2) non proliferazione, disarmo e sicurezza nucleare; 3) armi delle distruzioni di massa (WMD); 4) armi leggere ed armi leggere (SALW); e 5) contro-stupefacente. Le disposizioni sulla cooperazione riguardano i seguenti settori: sviluppo dell'infrastruttura, energia, trasporto, salute, risorse naturali, tassa, formazione e cultura, occupazione ed affari sociali, scienza e tecnologia ed ambiente e mutamento climatico. L'accordo inoltre sottolinea l'importanza della cooperazione legale ed afferma impegno dei partiti' alla criminalità organizzata, al riciclaggio di denaro ed alla corruzione combattenti.

2. RISULTATI DELLE CONSULTAZIONI CON LE PARTI INTERESSATE E LE VALUTAZIONI DI IMPATTO

Il Consiglio è stato tenuto informato in tutte le fasi dei negoziati. È stato consultato regolarmente nel gruppo di lavoro competente sull'Asia.

Il Parlamento Europeo inoltre è stato tenuto completamente informato in tutto i negoziati e l'alto rappresentante dell'unione per gli affari esteri e la politica della sicurezza (" alto

¹ Foreign Affairs Council Conclusions of 18 July 2011 (doc. st12865/11)

² Council Decisions of 10 November 2011 (doc. st 16146/11 and st16147/11)

rappresentante ") inviati ad una copia dell'accordo siglato al presidente del Parlamento Europeo su.... ..

L'alto rappresentante e la Commissione considerano che gli obiettivi definiti dal Consiglio nelle sue direttive di negoziazione siano stati conseguiti e che il progetto d'accordo può essere presentato per la firma.

3. ELEMENTI LEGALI DELLA PROPOSTA

Base giuridica

Secondo giurisprudenza sistemata, la scelta della base giuridica deve riposare sui fattori obiettivi favorevoli a controllo giurisdizionale, che comprendono lo scopo ed il contenuto di quella misura.

Lo scopo dell'accordo, come descritto nella *natura dell'articolo 2 e nella portata*, è di stabilire un'associazione fra i partiti che riposa su un approccio multipillar, per rinforzare il dialogo e la cooperazione.

L'accordo riguarda la cooperazione politica (titolo II), cooperazione allo sviluppo (titolo III), cooperazione su commercio ed investimento (titolo IV), cooperazione negli argomenti di libertà, sicurezza e giustizia (titolo V), cooperazione settoriale (titolo VI) e cooperazione regionale (titolo VII). Inoltre, l'accordo comprende le disposizioni sulla messa a punto istituzionale (titolo VIII) e le disposizioni finali (titolo IX).

Un'analisi dettagliata dello scopo e del contenuto dell'accordo indica che alcune delle disposizioni della caduta di accordo nell'ambito della politica estera e di sicurezza comune ed alcuni altri elementi rientrano nella politica commerciale come pure nell'ambito della politica di cooperazione allo sviluppo dell'Unione Europea. L'analisi avanza indica che nessuno di questi elementi possono essere considerati come l'accessorio a c'è ne dell'altro ha citato gli elementi e nessuno dei tre elementi citati possono essere identificati chiaramente come la componente principale. Di conseguenza, la proposta dovrebbe essere basata sulle basi giuridiche multiple, vale a dire l'articolo 37 TUE e TFEU degli articoli 207 e 209.

Natura giuridica

Le basi giuridiche si sono riferite a sopra confer potere sull'UE di prendparteeere agli accordi con i terzi nei settori che l'accordo tratta. In particolare:

- L'articolo 37 TUE assegna all'UE il potere di concludere gli accordi con i terzi nell'area di politica estera e di sicurezza comune.
- Concessioni dell'articolo 207 TFEU l'UE il potere di concludere gli accordi nell'area di politica commerciale comune. L'articolo 3(1) TFEU dà all'UE la competenza esclusiva in questa area.
- Concessioni dell'articolo 209 TFEU l'UE il potere di concludere gli accordi con i paesi extracomunitari nell'area di politica di cooperazione allo sviluppo.

Conformemente a giurisprudenza, alla domanda se una misura particolare da un accordo *rientra nella competenza del [dell'unione] è una che si riferisce all'attribuzione e, così, all'esistenza stessa di quella competenza e non alla sua esclusiva o natura comune*'. Di conseguenza e nella misura in cui i Trattati hanno attribuito la competenza all'unione in un'area di politica, l'unione può esercitare tale competenza esternamente.

Conformemente all'articolo 3(2) TFEU, l'UE inoltre ha competenza esclusiva per la conclusione di un `di accordo internazionale in quanto la sua conclusione può colpire le regole comuni o alterare la loro portata'. La Corte di Giustizia ha sostenuto in una decisione

che i negoziati possono colpire le regole comuni o alterare la loro portata nella misura in cui il contenuto dei negoziati [...] le cadute del ` all'interno di un settore trattato in larga misura dall'UE comune governa'... Di conseguenza quei negoziati fanno parte della competenza esclusiva dell'Unione Europea,, '. La giurisprudenza recente inoltre chiarisce che tutta l'analisi di una competenza, particolarmente una competenza esclusiva, deve considerare non solo le regole attuali del sindacato ma anche lo sviluppo di futuro prevedibile del ` [...] di quelle regole,, '.

I Trattati hanno potere così conferito sull'UE di agire in tutti i settori trattati dall'accordo.

Altre considerazioni di ordine morale

L'accordo ha installato un quadro istituzionale composto di comitato misto (vedi il *quadro istituzionale* di titolo VIII, dell'articolo 49). Il comitato misto può installare i comitati speciali o i gruppi di lavoro per assisterlo con la prestazione delle sue mansioni. *L'adempimento dell'articolo 54 degli obblighi* inoltre ha precisato una procedura per la sistemazione delle dispute, per essere usato se uno dei partiti non riesce ad ottemperare ai suoi obblighi in virtù dell'accordo.

L'accordo è valido per un periodo iniziale di dieci anni. Sarà automaticamente esteso per i periodi consecutivi successivi di cinque anni a meno che, sei mesi prima della scadenza della sua validità, uno dei partiti informi l'altro per iscritto della sua intenzione di non estenderlo. L'accordo può essere terminato a sei mesi' di avviso.

Proposta unita di a

DECISIONE DI CONSIGLIO

sulla firma, a nome dell'Unione Europea, dell'accordo di cooperazione sull'associazione e dello sviluppo fra l'Unione Europea e la Repubblica islamica di Afghanistan

IL CONSIGLIO DELL'UNIONE EUROPEA,

Considerando il trattato sull'Unione europea, in particolare l'articolo 37 di ciò,

Considerando il Trattato sul funzionamento dell'Unione Europea, in particolare degli articoli 207 e 209, insieme con l'articolo 218(5) ed il secondo comma dell'articolo 218(8) di ciò,

Vista la proposta della Commissione Europea e dell'alto rappresentante dell'unione per gli affari esteri e la politica della sicurezza,

Considerando che:

- (1) nel novembre 2011, il Consiglio ha autorizzato la Commissione e l'alto rappresentante dell'unione affinché gli affari esteri e la politica della sicurezza avvii i negoziati con la Repubblica islamica di Afghanistan su un accordo di cooperazione sull'associazione e sullo sviluppo³.
- (2) i negoziati sull'accordo di cooperazione riuscivano e l'accordo è stato siglato il 2 luglio 2015 in Kabul.
- (3) l'accordo dovrebbe quindi essere firmato a nome dell'UE salvo sua conclusione ad una data ulteriore,

HA ADOTTATO QUESTA DECISIONE:

L'articolo 1

La firma dell'accordo di cooperazione sull'associazione e dello sviluppo fra l'Unione Europea e la Repubblica islamica di Afghanistan è approvata con ciò a nome dell'unione salvo conclusione dell'accordo detto.

Il testo dell'accordo è attaccato a questa decisione.

L'articolo 2

Il segretariato generale del Consiglio stabilirà lo strumento delle piene potenze firmare l'accordo salvo sua conclusione, dato che le persone indicate dai negoziatori dell'accordo.

L'articolo 3

Questa decisione entrerà in vigore il giorno che segue quello della sua approvazione.

³ Council Decisions of 10 November 2011 (doc. st 16146/11 and st16147/11)

Fatto a Bruxelles,

*Per il Consiglio
Il presidente*