



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA

## Università degli Studi di Padova

Dipartimento di Studi Linguistici e Letterari

Corso di Laurea Magistrale in  
Lingue Moderne per la Comunicazione e la Cooperazione Internazionale  
Classe LM-38

Tesi di Laurea

### *Il post-editing: un'analisi della sua percezione e della compensazione equa*

Relatrice  
Prof.ssa Sabina Fata

Laureanda  
Greta Maniero  
n° matr. 2027725 / LMLCC

Anno Accademico 2021 / 2022



## Indice

<b>Abstract</b>	p.	3
<b>Introduzione</b>	p.	5
<b>Capitolo 1: Machine Translation</b>	p.	11
1.1 La Traduzione Automatica	p.	11
1.2 Storia della traduzione automatica	p.	14
1.2.1 Primi pionieri: 1930-1954	p.	14
1.2.2 Tra innovazione e disillusione: 1954-1965	p.	15
1.2.3 Il rapporto ALPAC e il suo impatto: 1966-1975	p.	16
1.2.4 Il grande sviluppo della MT: 1976-1996	p.	18
1.2.5 Dai primi sistemi gratuiti online al giorno d’oggi: 1997-2023	p.	19
1.3 I sistemi di traduzione automatica	p.	21
1.3.1 Rule-Based Machine Translation	p.	22
1.3.2 Corpus-based Machine Translation	p.	26
1.3.3 Statistical Machine Translation	p.	28
1.3.4 Neural Machine Translation	p.	32
1.4 Traduzione Assistita	p.	37
1.4.1 Computer-Aided Translation	p.	37
1.4.2 Translation Memory Systems	p.	39
<b>Capitolo 2: Post-editing</b>	p.	45
2.1 Cosa si intende con “post-editing”	p.	45
2.1.1 Post-editing e revisione a confronto	p.	46
2.1.2 ISO 18587 – Post-Editing Standard	p.	49
2.2 Tipi di post-editing	p.	52
2.2.1 Light PE	p.	52
2.2.2 Full PE	p.	53
2.2.3 PE monolingue e bilingue	p.	55
2.3 Competenze	p.	57
2.3.1 Competenze dei post-editor	p.	58
2.3.2 Formazione dei post-editor	p.	61
2.4 Valutazione del post-editing	p.	64
2.4.1 Errori tipici dei sistemi di traduzione automatica neurale	p.	64
2.4.2 Produttività e qualità	p.	66

2.5 Post-editing nei translation memory systems	p.	69
2.6 Rischi e riservatezza dei dati nel post-editing	p.	72
2.6.1 Rischi del post-editing	p.	73
2.6.2 Protezione dei dati	p.	75
2.7 Prospettive future	p.	77
2.8 Fattori di (non) adozione del post-editing	p.	80
<b>Capitolo 3: Indagine sulla compensazione equa nel post-editing</b>	p.	85
3.1 Introduzione	p.	85
3.2 Metodi di pagamento del post-editing	p.	87
3.2.1 Pagamento a parola	p.	89
3.2.2 Pagamento su base oraria	p.	90
3.2.3 Pagamento edit-distance	p.	91
3.2.4 Direttive della società TAUS sulla compensazione	p.	94
3.3 Metodologia: associazione libera	p.	95
3.4 Prima sezione: percezione del post-editing e della sua retribuzione	p.	99
3.4.1 Compito di associazione libera	p.	100
3.4.2 Analisi dei risultati	p.	108
3.5 Seconda sezione: il punto di vista dei post-editor	p.	116
3.5.1 Analisi dei risultati	p.	116
3.5.2 Considerazioni finali fornite dai post-editor	p.	127
<b>Conclusione</b>	p.	133
<b>Appendice</b>	p.	139
<b>Bibliografia</b>	p.	147
<b>Sitografia</b>	p.	151
<b>Zusammenfassung</b>	p.	153

## Abstract

Nel seguente progetto di tesi viene trattata la compensazione equa del post-editing. In particolare, obiettivo della dissertazione è quello di comprendere quale sia la percezione del post-editing non solo tra i post-editor stessi, ma anche tra il pubblico generale, attraverso il metodo dell'associazione libera: sono state sottoposte due parole stimolo, "translation" e "post-editing", alle quali i partecipanti hanno associato dai 2 ai 5 nomi o aggettivi. È stato poi chiesto ai post-editor quale fosse, secondo la loro opinione, il metodo di remunerazione più equo per chi si occupa di tale pratica, al fine di comprendere quale venga maggiormente incontro alle esigenze dei post-editor e sia proporzionale allo sforzo richiesto. Inizialmente, viene inoltre proposta un'analisi della letteratura presente riguardo alla *machine translation*, la sua storia e diversi sistemi di traduzione automatica esistenti; successivamente, viene descritto dettagliatamente cosa si intende con post-editing, le sue caratteristiche e difficoltà, giungendo, infine, al tema chiave della tesi, ovvero i metodi di pagamento e la compensazione equa.



## Introduzione

Il lavoro di tesi di seguito presentato ha per oggetto il post-editing e la sua compensazione. Con il termine “post-editing” si intende la revisione dell’output prodotto da un sistema di traduzione automatica da parte di un traduttore umano, accuratamente formato in quest’ambito, che assume il titolo di “post-editor”. Essendo un’attività distinta dalla traduzione, il post-editing prevede delle modalità diverse di remunerazione. È attualmente in corso un dibattito su tale tematica, non essendo ancora stato definito un metodo di pagamento unico. L’obiettivo principale dell’elaborato, dunque, è quello di indagare la percezione del post-editing sia tra i post-editor stessi che tra il pubblico generale, composto da studenti, imprenditori, impiegati, lavoratori dipendenti, ecc., i quali possono rappresentare dei potenziali clienti oppure potrebbero entrare in contatto con il post-editing o con la traduzione automatica nello svolgimento dei propri impieghi. Il fine, dunque, è quello di comprendere la loro opinione sull’argomento e possibilmente individuare il metodo di pagamento ritenuto migliore per il post-editing: con “migliore” è da intendersi un pagamento che si dimostra equo, ovvero che tiene conto delle diverse variabili previste dal post-editing, quali la qualità dell’output prodotto dal sistema di traduzione automatica, la coppia linguistica presa in considerazione, il contenuto del testo e il tempo necessario alla lettura del testo sorgente e del lavoro finale.

Alla base della scelta dell’oggetto di studio vi è l’interesse nei confronti del mondo della traduzione, in linea con il percorso universitario compiuto, e la curiosità verso un argomento così dibattuto, a cui ancora oggi non si è trovata una soluzione valida. Ho trovato personalmente molto interessante analizzare le risposte date dai post-editor e dal pubblico generale sull’argomento. Di particolare importanza si sono rivelati i commenti approfonditi rilasciati dai post-editor nella sezione finale del questionario proposto, che hanno fornito degli spunti di riflessione importanti.

L’elaborato è suddiviso in tre capitoli, il primo dei quali è dedicato alla traduzione automatica (*machine translation* - MT), una tipologia di traduzione prodotta da un computer, che coinvolge, in alcune circostanze, l’intervento umano, attraverso un lavoro di pre- o post-editing. In particolare, verranno analizzate le principali tappe che hanno

segnato la storia della traduzione automatica, dalle prime sperimentazioni e dai primi brevetti, presentati dai pionieri nella prima metà degli anni '30 del Novecento, quando la MT veniva ancora chiamata “traduzione meccanica”, fino all’evoluzione raggiunta ai nostri giorni. Oggigiorno, infatti, è possibile affermare che la traduzione automatica è divenuta un bene di consumo di massa, e chiunque può usufruire dei sistemi MT gratuiti e utilizzabili online. Gran parte degli eventi storici nominati sono tratti dagli scritti di Hutchins (1995; 2000), in cui sono riportati i problemi iniziali e le sfide affrontate nel corso dell’evoluzione della traduzione automatica.

Verranno, poi, analizzati nel dettaglio i principali sistemi esistenti, che si sono susseguiti nel corso di più di 70 anni di storia: la traduzione automatica basata su regole, ovvero *rule-based machine translation* (RBMT), dominante tra gli anni '50 e '80 del secolo scorso; la traduzione automatica statistica, o *statistical machine translation* (SMT), dagli anni '90 fino al 2018; infine, la traduzione automatica neurale, *neural machine translation* (NMT), che ha preso piede dal 2018 ed è considerata quella maggiormente all’avanguardia. Saranno analizzate le principali somiglianze e differenze esistenti tra i tre sistemi: in modo particolare, il sistema RBMT è basato su regole, mentre i sistemi SMT e NMT, essendo *data-driven*, ossia basati su dati, si servono dei risultati dell’analisi dei dati di addestramento; i tre sistemi differiscono, infatti, sia nel loro funzionamento che nella loro architettura.

Verrà indicato cosa sono e come funzionano i sistemi di “traduzione assistita” (*computer-aided translation* – CAT), nata negli anni '60 del Novecento con il fine di aiutare i traduttori (umani) nello svolgimento del proprio lavoro, divenendo oggigiorno uno strumento indispensabile nel lavoro quotidiano del traduttore. In modo particolare, verrà spiegato, per l’appunto, il loro funzionamento, basato sulle memorie di traduzione (*translation memory*), la loro composizione e architettura, evidenziando le principali differenze esistenti con la traduzione automatica.

Nel corso del secondo capitolo, invece, verranno affrontate nel dettaglio le diverse sfaccettature del post-editing della traduzione automatica. Innanzitutto, verranno esaminate le differenze e le somiglianze esistenti tra post-editing e revisione, che, infatti, sono da considerarsi come dei compiti tra loro diversi. Il post-editing della MT è regolato da uno standard internazionale, denominato ISO 18587 e volto a fornire i

requisiti per il processo di post-editing. Inoltre, verranno elencate le linee guida di tale attività: è importante sottolineare che vi sono delle linee guida da seguire, per un corretto svolgimento del lavoro di post-editing, fornite dalla società TAUS (Translation Automation User Society). Esistono, infatti, due gradi di post-editing in base alle modifiche da applicare, ovvero *light post-editing* e *full post-editing*, richiesti dal cliente e a cui il traduttore si deve attenere, affinché le tempistiche previste per lo svolgimento del lavoro vengano rispettate. Inoltre, il post-editing normalmente è bilingue, ovvero viene confrontato il testo di partenza, nella lingua originale, ma, in alcuni casi può anche essere effettuato senza il suo impiego, diventando monolingue.

Verrà poi proposto un focus sugli errori tipici commessi dai sistemi di traduzione automatica neurale: questi sistemi, sebbene siano molto più precisi rispetto ai precedenti sistemi statistici, commettono degli errori che normalmente non verrebbero compiuti da un traduttore umano. Infatti, sebbene l'output risulti fluido, vi sono spesso delle omissioni o altri errori, che non permettono di definirlo perfetto. Collegato alla presenza o meno di errori nell'output, verrà spiegato cosa si intende per produttività nel post-editing, ovvero il numero di parole di cui il traduttore è in grado di svolgere il post-editing nel corso di un giorno. Tuttavia, la produttività dipende da una serie di fattori come la combinazione linguistica, il tipo di contenuto e il supporto tecnologico a disposizione del post-editor. Inoltre, la produttività è influenzata anche dalla qualità dell'output prodotto dal sistema di traduzione automatica: più la qualità è buona, più sono le probabilità di aumentare la propria produttività; viceversa, se la qualità dell'output della MT risulta molto scarsa, le modifiche da apportare saranno in numero maggiore, provocando un rallentamento e una diminuzione della produttività.

Verranno inoltre elencate le diverse conoscenze che il post-editor deve possedere per svolgere il proprio lavoro nel migliore dei modi, accanto a una corretta formazione e educazione nell'ambito. Il post-editor deve essere anche consapevole dei rischi che si incorrono nello svolgimento di tale impiego. Infatti, la traduzione è un lavoro che deve essere svolto in maniera precisa, sotto tutti i suoi aspetti, perché può essa stessa essere considerata come un processo "rischioso", potendo incorrere in conseguenze importanti a livello legale. Particolare attenzione, infatti, va prestata alla protezione dei dati di clienti e utenti, inseriti nei sistemi MT: per questo l'analisi della politica sul trattamento dei dati personali è oggetto di studio dell'elaborato in questione, insieme ai rischi legati

ai possibili danni arrecati al cliente, rischi informatici e problemi di responsabilità, oltre a danni di reputazione del post-editor stesso, nel caso in cui il lavoro prodotto e consegnato presenti degli errori.

Una sezione è stata dedicata al futuro del post-editing: da un lato vi è un comune sentimento di paura da parte dei traduttori, che temono di essere messi in disparte e surclassati dalla traduzione automatica, ritrovandosi a dover effettuare solamente post-editing; dall'altro, invece, vi è la speranza di un continuo miglioramento della *machine translation*, con il fine di raggiungere la cosiddetta *singularity*, ottenuta quando un traduttore professionista umano impiega meno tempo per il post-editing della MT che non per la revisione di una traduzione (umana). Infine, sono state analizzate le principali motivazioni alla base della scelta dei traduttori di occuparsi di post-editing o meno: come verrà sottolineato, infatti, vi sono molti traduttori che tuttora sono fortemente contrari al post-editing e si rifiutano di occuparsene, nonostante l'aumento di richiesta di tale impiego da parte del mercato della traduzione e dei clienti.

Nel terzo capitolo, ovvero il capitolo finale, si entra nel vivo del tema della tesi, ossia la compensazione equa nel post-editing. Attualmente, non esiste una metodologia di remunerazione che sia valida e utilizzata da tutti i post-editor e i clienti: è proprio a quest'ultimi che, il più delle volte, spetta la scelta del metodo di pagamento da applicare. Le tipologie di remunerazione esistenti sono varie; tuttavia, le più utilizzate e conosciute sono il pagamento a parola, quello su base oraria e quello basato sulla distanza di modifica (*edit-distance*). Ognuno dei metodi citati possiede dei punti a favore e altri a sfavore, che verranno elencati nel corso del capitolo in questione. Come accennato, da anni è in corso una discussione su quale tra questi metodi possa essere considerato il più equo, alla quale, al momento attuale, non si ha una soluzione definitiva. Gran parte dei post-editor viene normalmente pagata a parola, metodologia di norma utilizzata nella traduzione. Tuttavia, questo sistema di remunerazione non è sempre giudicato adatto al post-editing, poiché non permette di tenere conto delle variabili sopracitate; al contrario, il pagamento su base oraria, da questo punto di vista, risulterebbe maggiormente equo. Scopo del capitolo in questione è, infatti, quello di rappresentare le idee dei post-editor sull'argomento. Per farlo, è stato sottoposto un questionario a 132 partecipanti, metà dei quali appartenenti al pubblico generale, mentre

l'altra metà composta dai post-editor. Come si vedrà, il questionario è diviso principalmente in due sezioni: una prima parte, sottoposta all'attenzione di tutti i 132 partecipanti, volta allo studio della percezione del post-editing e della sua remunerazione, e una seconda sezione, rivolta solamente ai post-editor, per verificare quale, secondo loro, sia il metodo di remunerazione migliore. Al fine di comprendere al meglio le percezioni dei partecipanti, nella prima parte del questionario è stato richiesto di completare un compito di associazione libera, sottoponendo alla loro attenzione due parole-stimolo, *translation* e *post-editing*, a partire dalle quali i partecipanti hanno dovuto produrre dai due ai cinque sostantivi o aggettivi a esse collegati. Successivamente, sono state proposte una serie di domande chiuse e aperte, riguardanti la conoscenza del post-editing e dei suoi metodi di remunerazione. Nella seconda parte, invece, sono state presentate una serie di domande chiuse, a risposta singola o multipla, volte alla comprensione del grado di soddisfazione dei post-editor riguardo al metodo di pagamento normalmente ricevuto e a quello che, secondo loro, potrebbe essere, invece il pagamento più equo.

Concludendo, come anticipato, è stato dedicato uno spazio a eventuali riflessioni e pensieri aggiuntivi riguardo alla tema dell'elaborato, raccogliendo diverse testimonianze fornite da traduttori che si occupano di post-editing, analizzate nel corso del capitolo.

L'analisi proposta dall'elaborato in questione rappresenta, seppure in scala ridotta, il pensiero generale dei post-editor sul tema della remunerazione, permettendo una comprensione più approfondita delle problematiche a essa relative. Inoltre, racchiude la speranza di ottenere finalmente un maggiore riconoscimento del lavoro svolto e dello sforzo cognitivo richiesto ai post-editor, oltre a una migliore informazione sull'argomento da parte dei clienti, a cui normalmente è affidata la decisione del sistema di pagamento da utilizzare.



## I. Machine Translation

### 1.1 La Traduzione Automatica

Il termine *Machine Translation* (MT), o “Traduzione Automatica” in italiano, indica una traduzione svolta principalmente da un computer (inteso come un sistema di traduzione), coinvolgendo, in alcuni casi, l’intervento umano attraverso il pre- o post-editing (Bowker & Fisher, 2010:60). Il nucleo dell’idea della *machine translation* si basa sull’automazione dell’intero processo traduttivo (Hutchins, 1995:431). La traduzione automatica prevede, come la traduzione umana, la resa del testo di partenza in una determinata lingua di arrivo (*target language*); tuttavia, la produzione del testo nella lingua d’arrivo avviene automaticamente, tramite l’utilizzo di un sistema di traduzione, per l’appunto (Kenny, 2022:32). La traduzione automatica viene considerata come una delle prime applicazioni non numeriche dei computer sviluppati dopo la Seconda Guerra Mondiale (ibidem).

I primi approcci all’automatizzazione della traduzione possono essere considerati “primitivi”, se comparati agli standard odierni (ibidem). Ciononostante, oggi si dispone di un numero elevato di risorse e di una tecnologia altamente avanzata, dunque bisogna riconoscere gli sforzi e i risultati raggiunti fin dagli anni ’50 e ’60 dai primi ricercatori (Hutchins, 2000, in Kenny, 2022:32). I pionieri della traduzione automatica provenivano da contesti differenti: linguistica, ingegneria elettronica, fisica, filosofia, traduzione e interpretariato (Hutchins, 2000:1). Diverse furono le motivazioni alla base dell’ideazione di questa nuova tecnologia: secondo alcuni studiosi vi era la necessità di creare un sistema in grado di produrre, a sua volta, una traduzione “utile”, sebbene carente in accuratezza e stile; per altri si trattava di una sfida intellettuale, con l’obiettivo di rivelare il funzionamento del linguaggio e della mente umana; per altri ancora era vista come un importante banco di prova per la linguistica strutturale e formale (ibidem).

Tuttavia, i pionieri si dovettero confrontare fin dal principio con una serie di problemi, alcuni alquanto complessi. Infatti, sussistono delle differenze significative nel modo, tra le diverse lingue, di esprimere gli stessi concetti, idee, oggetti: le prime difficoltà riguardarono proprio la resa di queste differenze attraverso la traduzione automatica (ibidem). Se alcuni studiosi ritenevano che fosse sufficiente un lavoro di ampliamento di dizionari (dunque la creazione di sistemi basati su dizionari), relegando

a un secondo piano le differenze di struttura sintattica, altri ricercatori sostenevano che era essenziale “esplorare” la semantica interna della mente umana, indagando i processi di pensiero alla base della comunicazione (ibidem). La creazione di sistemi basati su dizionari si rivelò limitante, portando l’attenzione sull’importanza dell’analisi e della trasformazione delle strutture sintattiche: infatti, a partire dagli anni ’60 la sintassi divenne l’obiettivo principale dei gruppi di studiosi della MT (ibidem).

Un ulteriore grosso problema per i pionieri, durato a lungo, era costituito proprio dai computer stessi: erano limitati, sia per quanto riguarda la memoria interna che la velocità, costosi da utilizzare e non disponibili su larga scala (ivi, 3). I primi gruppi di ricercatori si dovettero focalizzare solamente su “simulazioni” manuali e utilizzare schede perforate, per arrivare alla prova su computer solamente in seguito (ibidem). In modo particolare, ciò riguardò alcuni gruppi di ricercatori russi, per i quali i computer non furono disponibili fino agli anni ’70 (epoca in cui i computer che divennero disponibili in Russia non erano comunque considerati equiparabili a quelli americani, in termini di capacità e velocità) (ibidem). Per questa ragione si parlava, durante i primi anni di sviluppo di tale tecnologia, di “traduzione meccanica”, termine che si riferiva all’inserimento macchinoso dei testi, che dovevano essere codificati su schede perforate (ibidem). L’inserimento dei primi lettori ottici di caratteri avvenne solamente negli anni ’80; inoltre non vi era un sistema di codifica univoco e la maggioranza dei gruppi di ricerca arrivò a ideare il proprio sistema di codifica (ibidem). L’output della traduzione, inizialmente, era costituito da risme di larghi fogli di carta per computer, stampati riga per riga in lettere maiuscole, spesso quasi illeggibili (ibidem). L’archiviazione *off-line* avveniva o su schede perforate, o su carta, o su nastri d’acciaio magnetizzati: tutti questi fattori contribuirono a rendere ancora più macchinosa questa tipologia di traduzione, che avrebbe dovuto, al contrario, agevolare il lavoro umano (o addirittura sostituirlo) (ibidem).

I pionieri della traduzione automatica riponevano in essa estrema fiducia, supponendo che fosse adatta alla traduzione di qualsiasi tipo di testo, eppure, ben presto si capì che i testi più adatti alla traduzione automatica erano quelli di tipo tecnico-scientifico (ibidem). Inoltre, si riteneva di poter ottenere, tramite la MT, una traduzione indistinguibile da quella umana o, perlomeno, equiparabile a essa. I ricercatori capirono, tuttavia, che sarebbe stato impossibile produrre una traduzione perfetta, ed

evidenziarono la necessità di un maggiore coinvolgimento umano sia nella revisione dell'output prodotto dal sistema di traduzione automatica (ovvero grazie al post-editing), che nel controllo dei testi all'ingresso (con il pre-editing), al fine di ottenere un output di maggiore qualità (ibidem).

I sistemi di traduzione automatica, inizialmente utilizzati solamente da governi e organizzazioni internazionali, iniziarono a essere adoperati anche dal pubblico generale, in modo particolare con l'avvento di Internet e il libero accesso ai primi sistemi gratuiti online (come Babel Fish di AltaVista): si può, infatti, citare il fenomeno Google Translate, che può essere considerato il sistema di traduzione automatica online gratuito più conosciuto, e che, nel 2016, aveva già raggiunto oltre il mezzo miliardo di utenti, traducendo più di 100 miliardi di parole al giorno in 103 lingue (Turovsky, 2016, in Kenny, 2022:32).

Al giorno d'oggi la MT è presente in diverse attività quotidiane: ad esempio, se in combinazione con dei motori di ricerca come Google o Microsoft Bing, è possibile ottenere la traduzione automatica di una pagina web straniera nella lingua dell'utente (ivi, 32-33). Inoltre, viene utilizzata insieme ad altre tecnologie come il riconoscimento vocale automatico, la sintesi vocale, il riconoscimento ottico dei caratteri, l'elaborazione digitale di immagini: in questo modo gli utilizzatori di tali tecnologie sono in grado di tenere conversazioni con persone che non parlano la propria lingua o leggere indicazioni stradali, cartelli, menù di luoghi in cui si parla una lingua diversa dalla propria, semplicemente servendosi del telefono (ivi, 33). La traduzione automatica è usata anche nella traduzione audiovisiva, nella traduzione di sottotitoli di film e serie tv: si può dunque affermare che la MT è divenuta un bene di consumo di massa (ibidem).

Nonostante l'enorme progresso nel campo della traduzione automatica, a cui si è potuto assistere nel corso degli ultimi anni, continuano a essere presenti, all'interno dell'output della traduzione, errori tra loro di diversa tipologia che, in determinati contesti quali quello sanitario o della diplomazia internazionale, possono portare a gravi conseguenze (ibidem). Vengono spesso sollevate delle polemiche riguardo ad altre questioni di tipo legale e morale, spesso in concomitanza con il post-editing, che verranno affrontate nel corso del capitolo **II. Post-Editing**.

Vengono di seguito analizzate le diverse fasi evolutive della traduzione automatica, dalla sua ideazione alla sua realizzazione, fino ai più recenti sviluppi tecnologici.

## 1.2 Storia della traduzione automatica

I seguenti paragrafi sono volti a riassumere i principali eventi nella storia della traduzione automatica, suddivisa in varie tappe, dal suo esordio fino al giorno d'oggi.

### 1.2.1 Primi pionieri: 1930-1954

Attorno agli anni '30 del secolo scorso si cominciò a parlare di *machine translation*, sebbene vi siano delle prove che fanno risalire l'idea della meccanizzazione dei processi traduttivi al XVIII secolo: si tratta delle cosiddette idee sui “caratteri reali” (*real characters*) e sulle “lingue universali” o “filosofiche” di Cartesio, Wilkins, Leibniz *et al.* (Hutchins, 2000:5). Tuttavia, due secoli più tardi, proprio nella prima metà degli anni '30, due ingegneri, il franco-armeno Georges Artsrouni, e il russo Pëtr Trojanskij, per primi sperimentarono e proposero dei brevetti per le *translating machines*, ovvero delle macchine in grado di tradurre (*ibidem*). Il contributo di Trojanskij fu particolarmente importante, in quanto egli propose sia un metodo per la formazione di un dizionario automatico bilingue, sia uno schema per la codifica dei ruoli grammaticali interlinguistici (basato sull'Esperanto), oltre a delle idee per l'analisi di frasi e la generazione di testi in altre lingue (Hutchins & Lovtskij, in Hutchins, 2000:5).

Tuttavia, furono Andrew Booth (London University) e Warren Weaver (Rockefeller Foundation) gli studiosi che nel 1946 e 1947 si incontrarono e proposero, con dei tentativi, delle prime idee per l'uso dei nuovi computer, indirizzati appositamente alla traduzione delle lingue naturali (*ibidem*). Nel 1948 Booth, insieme alla collaborazione di Richard H. Richens, lavorò a un sistema a schede perforate per una traduzione basata sul dizionario, integrando alcune analisi morfologiche (*ibidem*). All'epoca si parlava, come precedentemente accennato, di *mechanical translation*, ovvero “traduzione meccanica” (*ibidem*). Fu nel 1949, grazie alla pubblicazione di un memorandum da parte di Weaver, che venne illustrata e introdotta ufficialmente l'idea della *machine translation* al pubblico: le proposte specifiche di Weaver, basate sulle sue conoscenze di crittografia, statistica, logica, teoria dell'informazione e degli universali del linguaggio, segnarono un punto di partenza nella ricerca sulla traduzione automatica (*ibidem*).

Nel 1951 il filosofo-logista israeliano Yehoshua Bar-Hillel venne nominato “primo ricercatore nell'ambito della machine translation”: Bar-Hillel fornì, infatti, alcuni approcci basilari alle problematiche della traduzione automatica (*ivi*, 6). Un importante avvenimento da ricordare, avvenuto l'anno seguente (1952), fu la prima conferenza

sulla *machine translation*, tenuta al MIT (Massachusetts Institute of Technology), durante la quale, per la prima volta, vennero trattati apertamente argomenti quali il pre- e post-editing, i microglossari, l'analisi sintattica, suggerimenti riguardo alle future attività, nonché fu sottolineata da Léon Dostert la necessità di raccogliere finanziamenti per la ricerca, una volta provata la fattibilità della traduzione automatica (ibidem). Nel 1954, dunque, fu presentata presso la sede IBM (International Business Machines) di New York, frutto di una collaborazione tra Paul Garvin della Georgetown University e Peter Sheridan dell'IBM, una dimostrazione in scala ridotta di un sistema di traduzione automatica russo-inglese, denominato, per l'appunto "IBM-Georgetown system" (Hutchins, 1994, in Hutchins, 2000:6).

### **1.2.2 Tra innovazione e disillusione: 1954-1965**

La dimostrazione del sistema "IBM-Georgetown" suscitò l'interesse di non pochi studiosi nel campo, non solo negli Stati Uniti, ma in tutto il mondo: infatti, il russo Dimitrij Panov, dopo la dimostrazione del sistema IBM-Georgetown, avviò la ricerca sulla *machine translation* anche in Unione Sovietica, presso il suo Istituto di Meccanica di Precisione e Tecnologia Informatica (ITMiVT), dando inizio alla prima di una lunga serie di ricerche sulla traduzione automatica in questo paese (ibidem). Nel 1954 fu fondata al MIT, da William N. Locke e Victor Yngve, anche la prima rivista del settore: "Mechanical Translation" (ibidem). Da quel momento in poi, le attività di studio in questo ambito aumentarono, espandendosi con enorme rapidità, anche in alcune città europee, come a Milano, grazie a Silvio Ceccato, e presso l'Università di Cambridge, grazie a Margaret Mastermann; ciononostante, il primato nella ricerca continuò a essere tenuto da Stati Uniti e Unione Sovietica (ibidem).

In quegli anni di grande rivoluzione tecnologica si iniziò a parlare di intelligenza artificiale (considerata come una figlia della cibernetica), ufficialmente fondata nel settembre 1956 al Dartmouth College (ivi, 7). Infatti, gli anni '50 del secolo scorso, segnati dalla Guerra Fredda tra Stati Uniti e Unione Sovietica, interessarono una vasta gamma di esperimenti in diversi campi, che contribuirono a influenzare i ricercatori a fare sempre di più: va ricordato il lancio del primo satellite artificiale, sputnik, nell'ottobre 1957, un chiaro segnale nei confronti degli USA del progresso tecnico e tecnologico dell'URSS, spingendo le due macro-potenze a implementare gli studi anche nel campo della MT (ibidem).

Una tra le più importanti conferenze sulla traduzione automatica si tenne nel 1960, presso la University of California, Los Angeles (UCLA), in cui fu evidenziato un maggiore interesse verso l'analisi sintattica, sostituendosi alla precedente attenzione nei confronti degli approcci lessicografici (ivi, 8). Emersero, inoltre, alcuni dubbi riguardo all'effettiva possibilità di realizzare, in un futuro prossimo, dei sistemi di traduzione in grado di produrre un output di buona qualità, riconoscendo la semantica come il principale problema: le prospettive verso una FAHQT, ovvero una *full-automatic, high-quality translation*, non erano delle migliori e, in quegli anni, molti ricercatori arrivarono a pensare che non fosse fattibile, giungendo alla tragica conclusione che la MT fosse un completo fallimento, anche a causa del rapporto critico di Bar-Hillel sulla traduzione automatica (ivi, 9). Si ebbe, così, una disillusione generale, dovuta alla mancanza di progressi notevoli, di concretizzazione dei benefici attesi dall'analisi sintattica e a un mancato aumento della qualità dell'output prodotto (ibidem). Tuttavia, continuarono a essere fondati nuovi gruppi per lo studio della traduzione automatica (ad esempio a Nancy, in Francia, con Bernard Pottier, e alla Ohio State University, con S-Y. Wang, per lo studio della MT cinese), e continuarono a essere tenute numerose conferenze (ivi, 9-10).

I primi segnali di svolta si notarono finalmente nel 1963, con la richiesta, da parte dell'NSF (National Science Foundation) dell'istituzione di un comitato volto a indagare sullo stato della MT e a valutare se il sostegno finanziario da parte delle agenzie di finanziamento degli Stati Uniti dovesse continuare: nacque, dunque, nell'aprile 1964 l'ALPAC (Automatic Language Processing Advisory Committee) (ivi, 10-11). In quello stesso mese si tenne il primo seminario congiunto USA-Giappone, sulla traduzione automatica: fu un evento importante, in quanto in quell'occasione fu riconosciuto che la traduzione automatica non interessava solamente le lingue europee e il russo, ma aveva raggiunto anche i paesi orientali (ivi, 11).

### **1.2.3 Il rapporto ALPAC e il suo impatto: 1966-1975**

Nel novembre del 1966, l'ALPAC pubblicò un importante rapporto che invitava a interrompere il finanziamento per la ricerca sulla traduzione automatica negli Stati Uniti: tale provvedimento diede inizio a un effetto domino, portando al ritiro dei sostegni economici in altri paesi, come ad esempio in Russia (ibidem). La motivazione era chiara: non sussisteva, secondo l'ALPAC, alcuna possibilità di ottenere di una

FAHQT, e fu affermato, dunque, che non sussistesse alcuna speranza futura per la traduzione automatica (ibidem). Il rapporto era intitolato “Languages and machines: computers in translation and linguistics” e in esso fu presentata un’attenta analisi dei fondi impiegati nello sviluppo e nella ricerca della traduzione automatica, evidenziando i punti deboli e vacillanti di essa, che portarono, per l’appunto, al ritiro del sostegno economico (ALPAC, 1966).

Ciononostante, i ricercatori continuarono, seppure in modo meno visibile, a fare ricerca e, nel 1969, il tedesco Peter Toma cominciò a sviluppare, in collaborazione con l’IBM, il sistema “Systran”, il primo sistema di traduzione automatica statistica in commercio (Toma, 2000:136). Systran è tuttora una delle più vecchie aziende operanti nel campo della *machine translation* e inizialmente si occupava di traduzione dal russo all’inglese per conto dell’aeronautica militare degli Stati Uniti d’America (United States Air Force) (ibidem). Il contributo di Toma, fornendo le sue abilità sia a livello linguistico che di programmazione, fu essenziale (ivi, 144). Il successo fu dovuto alla notevolmente migliore qualità di Systran rispetto ai precedenti sistemi utilizzati: venne infatti sviluppato un linguaggio speciale per i linguisti, permettendo a quest’ultimi di “comunicare” direttamente con il computer, e riuscendo, in questo modo, a salvarsi dalla drastica diminuzione dei fondi, dovuta al rapporto ALPAC (ivi, 142). Systran sostituì il precedente traduttore di IBM, denominato Mark II (Hutchins, 2000:12). Tuttavia, i primi veri segnali pubblici della ribalta della traduzione automatica, dopo il rapporto ALPAC, si ottennero a metà degli anni ’70, con l’utilizzo, nel 1974, di Systran da parte della NASA per il lavoro di traduzione durante il progetto Apollo-Sojuz (ibidem). Successivamente, nel 1975, il funzionamento di Systran venne dimostrato a Bonn, Lussemburgo e Zurigo, e il sistema fu adottato dalla Commissione Europea come base per la traduzione automatica, conferendo, in tal modo, un’efficacia e un’utilità pratica (ibidem). Nell’anno successivo, inoltre, il gruppo TAUM (Traduction Automatique de l’Université de Montréal) sviluppò un nuovo sistema: Météo (ibidem). Météo fu il primo sistema di “sottolinguaggio” (*sublanguage system*) completamente automatico, utilizzato per la traduzione dall’inglese al francese delle previsioni del tempo (ibidem). Tuttavia, TAUM aveva anche un altro progetto, ovvero il “Q-system”, un metalinguaggio computazionale, per la manipolazione di alberi linguistici e stringhe,

e la fondazione del linguaggio di programmazione Prolog, largamente utilizzato nell'elaborazione della lingua naturale (Hutchins, 1995:436).

#### **1.2.4 Il grande sviluppo della MT: 1976-1996**

Dopo Systran e Météo, il mercato della traduzione automatica si espanse e i ricercatori lavorarono a una serie di nuovi sistemi: ad esempio, in Germania, a Saarbrücken, dalla prima metà degli anni '70 alla seconda metà degli anni '80, venne sviluppato SUSY (Saarbrücker Übersetzungssystem), un sistema di trasferimento multilingue altamente modulare (Nitzke & Hansen-Schirra, 2021:18). Successivamente, dal 1978 fu elaborato EUROTRA, un sistema MT, richiesto dalla Commissione Europea, in grado di tradurre le (all'epoca) nove lingue ufficiali della Comunità Europea (King & Perschke, 1982). Un altro sistema fu quello prodotto dalla Logos Corporation: l'azienda inizialmente ideò un sistema di traduzione inglese-vietnamita per la traduzione di manuali di velivoli durante gli anni '70; nel corso del decennio successivo, invece, si occupò della traduzione di altre coppie linguistiche, quali tedesco-inglese e inglese-francese (Hutchins, 1995: 437). Sistemi quali Systran e Logos vennero descritti da Hutchins (1995:437) come dei sistemi *mainframe*, ovvero sistemi centrali, progettati per un'applicazione generale, dal punto di vista teorico, mentre nella pratica i loro dizionari erano adattati solo a determinati domini tematici, come l'energia nucleare o l'ingegneria aerospaziale.

Il sistema MT che, tuttavia, era all'epoca considerato il più sofisticato, tra quelli in commercio, era METAL, sistema che traduceva la coppia linguistica tedesco-inglese, nato dalle ricerche, fin dalla fine degli anni '70, dall'azienda Siemens, a Monaco, disponibile al pubblico, però, solamente dal 1988 (ibidem). Negli anni '80 fu ideato anche uno dei primi sistemi di traduzione giapponesi (traduzione giapponese-inglese) denominato "Mu", sviluppato dall'Università di Kyoto, le cui caratteristiche principali erano l'uso dell'analisi dei casi grammaticali e la creazione di un ambiente di programmazione per la scrittura grammaticale (GRADE) (ivi, 438).

Lo sviluppo di metriche di valutazione della traduzione automatica iniziò contemporaneamente allo sviluppo di sistemi statistici di traduzione automatica: inizialmente le valutazioni furono effettuate da periti/"ispettori" umani, secondo il FEMTI (Framework for the Evaluation of Machine Translation in ISLE) e poi, tra il 1992 e il 1994 un'agenzia di ricerca dell'esercito degli USA, chiamata DARPA, si

occupò di esaminare l'output non editato dei sistemi MT, confrontando misure automatiche e giudizi umani sull'adeguatezza, fluidità e originalità/creatività (Nitzke & Hansen-Schirra, 2021:19).

L'uso della MT accelerò fortemente nel corso degli anni '90, e tale aumento si notò maggiormente in agenzie commerciali, servizi governativi e aziende multinazionali, in cui le traduzioni, principalmente di documenti tecnici, venivano prodotte (e attualmente vengono prodotte) in enorme scala (Hutchins, 1995:443). Prima degli anni '80, spesso si pensava che lo scopo della traduzione automatica fosse la sostituzione (per lo meno parziale) del traduttore umano (ibidem). Negli anni '90, invece, il focus si spostò sugli strumenti, volti ad agevolare il traduttore nel suo lavoro: si trattava, dunque, dell'accesso a dizionari e banche-dati terminologiche, dell'elaborazione di tesi multilingue, della gestione di glossari e risorse terminologiche, ecc. (ibidem). In quegli anni nacquero, per l'appunto, delle stazioni di lavoro per i traduttori, come ad esempio TRADOS (TRAnslation and DOcumentation Software). TRADOS fu il primo *translation memory system* in commercio, nato a Stuttgart, in Germania, dall'idea di Jochen Hummel e Iko Knyphausen nel 1984, un software che usava un sistema di corpora bilingue allineato (Nitzke & Hansen-Schirra, 2021:19, Hutchins, 1995:443).

### **1.2.5 Dai primi sistemi gratuiti online al giorno d'oggi: 1997-2023**

Nel 1997 fu lanciato il primo sistema gratuito di traduzione automatica online, ovvero Babel Fish su AltaVista (quello che poi diverrà Yahoo, era già al tempo un motore di ricerca americano), gratuito per tutti gli utilizzatori; il suo utilizzo fu interrotto, tuttavia, nel 2008 (Nitzke & Hansen-Schirra, 2021:20). In realtà, Systran offriva già un servizio online di traduzione di intere pagine web e Babel Fish nacque in seguito (ibidem). Tra il 2001 e il 2015 emersero diversi sistemi di valutazione automatica della MT, come BLEU (Bilingual Evaluation Understudy), che consisteva in una serie di misure statistiche, volte a valutare le similarità negli output prodotti da sistemi SMT e quelli frutto di traduzioni umane (ibidem); oppure NITS (National Institute of Standards and Technology), basato sugli stessi principi di BLEU, e METEOR (Metric for Evaluation of Translation with Explicit ORdering; Banerjee & Lavie, 2005, in Poibeau, 2017:91), risalente al 2005, quindi il più recente tra i tre, e quello con i risultati migliori, in quanto teneva maggiormente conto della semantica, identificando le parole semanticamente "piene" (ovvero nomi, verbi e aggettivi) (Poibeau, 2017:91).

Nel 2002 “Language Weaver”, fondata da Kevin Knight e Daniel Marcu, fu la prima azienda a mettere in commercio un approccio statistico alla traduzione automatica e al trattamento della lingua naturale (Nitzke & Hansen-Schirra, 2021:20).

Nei primi anni 2000, a causa dell’espansione dell’Unione Europea, le lingue comunitarie divennero venti: questo aumento rese necessari l’introduzione e l’utilizzo di sistemi che potessero aiutare i traduttori umani, o addirittura tradurre al loro posto: infatti, in alcuni tempi, l’Unione Europea arrivò a produrre più di 1.4 milioni di pagine di testo al mese, una quantità impensabile per soli traduttori umani (ibidem). Conseguentemente all’aumento di richieste di combinazioni linguistiche da tradurre, cominciarono a essere messi in commercio dei *toolkit* (strumenti di lavoro) *Open Source* per i sistemi di traduzione automatica: alcuni esempi sono GIZA ++ (uno strumento di allineamento per SMT), Moses (una piattaforma per creare sistemi SMT), Apertium (una piattaforma per la creazione di *Rule-Based* MT), Joshua (un decoder per SMT gerarchici, basati sulla sintassi) e META-SHARE (il database per i progetti dell’Unione Europea) (ivi, 21).

Come già sottolineato, dagli anni ’90 del secolo scorso cominciarono a essere disponibili diversi sistemi di traduzione automatica anche su Internet, molto spesso pure gratuitamente: al giorno d’oggi il sistema più conosciuto è sicuramente Google Translate. Inizialmente basato su Systran, Google sviluppò successivamente una propria tecnologia, prima per il russo, il cinese e l’arabo nel 2005, per poi creare un sistema online in grado di tradurre 25 coppie linguistiche nell’ottobre del 2007 (al giorno d’oggi, invece, gestisce più di 100 lingue, la cui qualità è molto variabile a seconda della coppia linguistica presa in considerazione) (Poibeau, 2017:99). Nel primo sistema, affinché potessero essere inseriti un grande numero di dati, vennero utilizzate le trascrizioni delle Nazioni Unite e del Parlamento Europeo; in seguito, venne utilizzato l’inglese come lingua di trasferimento (*transfer language*) (Nitzke & Hansen-Schirra, 2021:21). Nello stesso periodo nacque un nuovo progetto, il progetto Euromatrix, volto a procurare dei sistemi di traduzione automatica per tutte le lingue dell’Unione Europea, con un totale di circa 500 coppie linguistiche (ibidem).

Nel 2016 Google Translate introdusse la traduzione automatica neurale (NMT), sviluppando un sistema chiamato “Google Neural Machine Translation system” (GNMT), che consentì di migliorare l’accuratezza e la fluidità dell’output prodotto

(ibidem). Nello stesso anno SDL istituì “Adaptive MT” per SDL Trados Studio 2017, un motore di traduzione automatica di autoapprendimento (ibidem). Accanto a Google Translate, ai giorni nostri DeepL, un altro sistema NMT, introdotto dal dizionario online e fornitore di corpora Linguee, è divenuto uno dei sistemi online di traduzione automatica più conosciuto e ben funzionante (ovviamente con determinate coppie linguistiche) (ibidem).

Le evoluzioni del futuro sono ancora incerte, ma, sicuramente, la vicinanza alla perfezione sarà sempre più possibile con il corso del tempo e il miglioramento delle tecnologie.

### **1.3 I sistemi di traduzione automatica**

Esistono, principalmente, tre tipologie di sistemi di traduzione automatica, che si sono succeduti a partire dagli anni Cinquanta, solitamente suddivisi in base alle loro rispettive architetture computazionali (Melby, 2020:419): *rule-based machine translation* (RBMT), ovvero la traduzione automatica basata su regole; *statistical machine translation* (SMT), ossia la traduzione automatica statistica; *neural machine translation* (NMT), cioè la traduzione automatica neurale. I sistemi SMT e NMT si differenziano dai sistemi RBMT, in quanto, nel loro funzionamento, applicano i risultati dell’analisi dei dati di addestramento, effettuata automaticamente; tali dati consistono in vari testi di partenza (sorgenti) e nelle loro rispettive traduzioni, effettuate da traduttori umani (Melby, 2020:419). Per questa ragione, suddetti sistemi vengono definiti come *data-driven*, ovvero basati su dati anziché su regole, sebbene tutti dipendano da qualche tipo di dati (ibidem). Si potrebbe pensare che tutti i sistemi di traduzione automatica siano basati su regole, essendo essi stessi delle applicazioni informatiche che implementano una serie di regole da seguire, ovvero un algoritmo (ivi, 420). Diversamente dalle formule algebriche, ad esempio, le regole all’interno di sistemi basati su dati non sono esaminabili dall’uomo, non essendo simboliche, in modo particolare nei sistemi NMT (ibidem). Vi sono due categorie di regole, dunque, chiamate “simboliche” e “sub-simboliche”: le prime rappresentano quelle considerate ispezionabili, presenti nei sistemi RBMT, mentre le seconde sono quelle considerate “opache” e, quindi, non ispezionabili, all’interno dei sistemi NMT (Lieberman, 2016, in Melby, 2020:420).

I tre sistemi di traduzione automatica sopracitati si sono susseguiti nel corso dello scorso secolo e in quello corrente. Inizialmente, tra gli anni '50 e '80 i sistemi RBMT erano i predominanti; gli anni '90, invece, furono un periodo di importante transizione tra i sistemi RBMT e i sistemi SMT, che tuttavia, presero totalmente piede solamente a partire dai primi anni 2000, rimpiazzando completamente il precedente sistema basato su regole (Melby, 2020:420). Dal 2018 i sistemi *data-driven* NMT sono divenuti i predominanti, sebbene i vecchi paradigmi, insieme alla traduzione umana (che in alcuni casi risulta ancora la soluzione migliore), continuano a rappresentare una risorsa fondamentale, particolarmente in alcune combinazioni linguistiche sulle quali non si possiede una quantità sufficiente di dati (ibidem). Risulta interessante notare come i tre sistemi fossero già tutti esistenti fin al 1949, sebbene gli ultimi due avessero ancora bisogno di ulteriori sviluppi chiave, al di fuori del settore della traduzione, per diventare efficacemente funzionanti (ibidem).

Nel corso della seguente sezione verranno dunque esaminate le sostanziali differenze tra i diversi sistemi, evidenziando le particolarità di ognuno di essi.

### **1.3.1 Rule-Based Machine Translation**

I sistemi *Rule-Based*, come suggerisce la loro denominazione, si basano su delle regole, studiate appositamente dall'uomo (Melby, 2020:419). Infatti, la struttura dominante nella ricerca nel campo della MT negli anni '80 era basata essenzialmente su regole linguistiche di diverso tipo: regole per la suddivisione di una frase in parole singole; regole per la ricerca di parole all'interno di un dizionario; regole per l'analisi sintattica (basate su dati recuperati dal dizionario); regole per il trasferimento lessicale e sintattico; regole morfologiche; regole per la generazione del testo nella lingua di arrivo dalla rappresentazione intermedia (Melby, 2020: 419; Hutchins, 1995:440). METAL, SUSY, Mu, EUROTRA, erano tutti sistemi con approccio *rule-based*, dominanti negli anni '80 (ibidem).

Di particolare rilievo furono le ricerche di Vauquois nell'ambito della RBMT. Bernard Vauquois, un ricercatore francese nel campo della traduzione automatica, nel 1968 creò il cosiddetto "triangolo di Vauquois" (riprodotto nella Figura 1). Si tratta di un diagramma che racchiude in sé i principali approcci esistenti all'epoca nella *machine translation*, illustrando la traduzione automatica come un processo che a sua volta coinvolge tre principali processi: l'analisi del testo sorgente, il trasferimento da una

*source representation* a una *target representation*, ovvero il trasferimento di significato dal testo sorgente al testo di arrivo, e infine la generazione del testo di arrivo (Schwartz, 2018:171).

We can imagine as many levels as we wish from the zero level (level of the text considered as a string of characters), asymptotically towards a level of understanding. At each level a formalization of the input sentence can be defined. Then, it may be assumed that the deeper the level chosen, the easier the transfer is. At the limit, if the ideal level of understanding could be reached for a given sentence in one language, the same structural specifier would represent all the paraphrases of this sentence in all languages<sup>1</sup> (Vauquois, 1976, in Schwartz, 2018:171).

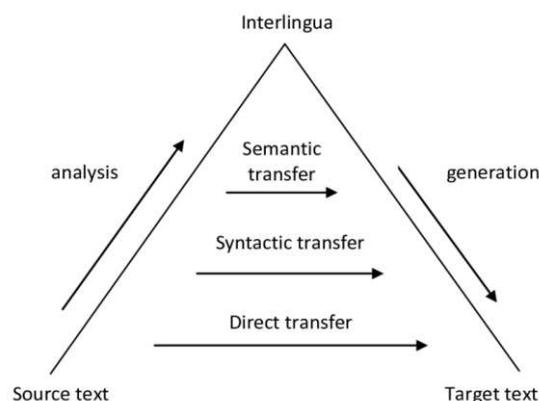


Figura 1: Triangolo di Vauquois

Vi sono tre principali approcci da tenere in considerazione quando si parla di rule-based machine translation, rappresentati all'interno del triangolo di Vauquois:

- *Direct transfer*: con il termine *direct transfer* si intende una traduzione diretta, da una lingua all'altra, senza una rappresentazione intermedia, differenziandosi da quella *indirect*, indiretta, in cui il sistema cerca, per prima cosa, di fornire una rappresentazione più astratta del contenuto che deve essere tradotto (Poibeau, 2017:21). La tipologia di approccio *direct transfer* è normalmente impostata sui dizionari, i quali consentono di ottenere una *word-for-word translation* basata, ovvero, sulla traduzione di ogni singola parola del testo sorgente nella corrispondente parola nella lingua di arrivo, applicando i cambiamenti morfologici necessari (ibidem). Successivamente, per ottenere un ordine di parole

---

<sup>1</sup> Possiamo immaginare tutti i livelli che desideriamo a partire dal livello zero (il livello del testo considerato come una stringa di caratteri), asintoticamente verso un livello di comprensione. A ogni livello si può definire una formalizzazione della frase in ingresso. Si può quindi ipotizzare che quanto più profondo è il livello scelto, tanto più facile è il trasferimento. Al limite, se il livello di comprensione ideale potesse essere raggiunto per una data frase in una lingua, lo stesso specificatore strutturale rappresenterebbe tutte le parafrasi di questa frase in tutte le lingue (tradotto da me).

il più vicino possibile a quello richiesto dalla lingua di arrivo, vengono utilizzate una serie di regole, più o meno sofisticate. Victor Oswald fece una serie di esperimenti sul trasferimento diretto e giunse alla conclusione che questo tipo di approccio di traduzione automatica non poteva funzionare (Schwartz, 2018:172-173). L'analisi sintattica in questo tipo di approccio, infatti, è poca o nulla (ibidem). Inoltre, ogni lingua possiede le sue specificità e risulta troppo banale ridurre la traduzione a una semplice trasposizione di singole parole da una lingua a un'altra. Può, tuttavia, contribuire a fornire un aiuto nel comprendere il significato generale del testo o captare qualche informazione sul suo contenuto e risultare accettabile nel caso di lingue tra di loro molto vicine a livello sintattico o di famiglia linguistica (Poibeau, 2017:22).

- *Analysis-transfer-generation*: questo approccio è più complicato del precedente in quanto è suddiviso in tre fasi distinte (Poibeau, 2017:21; Schwartz, 2018:173). La prima delle tre fasi consiste nell'analisi sintattica del testo sorgente, grazie alla quale il processo di traduzione riesce a sfruttare la struttura delle frasi del testo sorgente, evitando la limitazione imposta dalla traduzione diretta (parola per parola) (Poibeau, 2017:21). Da questa analisi viene prodotta una rappresentazione intermedia astratta, come un albero sintattico della lingua del testo sorgente, o una rappresentazione semantica (Schwartz, 2018:173). Nella seconda fase (*transfer*) tale rappresentazione è trasformata nell'equivalente rappresentazione intermedia della lingua di arrivo (ibidem). Nell'ultima fase, infine, il testo d'arrivo viene generato dalla rappresentazione intermedia della lingua di arrivo (ibidem). Il risultato dovrebbe essere maggiormente idiomatico, purché la componente sintattica fornisca informazioni accurate sulla lingua di partenza e su quella di arrivo (Poibeau, 2017:21).
- *Interlingua*: l'interlingua è il terzo approccio, su cui sono basati i sistemi più ambiziosi (Poibeau, 2017:21). Si propone di rimuovere la necessità di una fase di trasferimento, utilizzando una vera rappresentazione intermedia interlinguistica (Schwartz, 2018:174). "Interlingua" è il punto nel diagramma di Vauquois dove le rappresentazioni di entrambi i testi (*source* e *target*) diventano identiche, creando dunque una rappresentazione intermedia linguisticamente neutra (Melby, 2020:423). Ciò avviene a sua volta in due fasi: il testo sorgente viene sottoposto a

una profonda analisi, risultando in una rappresentazione intermedia indipendente linguisticamente e, per quest'ultima ragione, non vi è la necessità di una fase di trasferimento; la seconda fase genera direttamente il testo nella lingua di arrivo dalla rappresentazione intermedia (Schwartz, 2018:174). L'interlingua si diversifica da una *pivot language*, una lingua naturale che fa da ponte tra due lingue (ad esempio, per tradurre dal russo al giapponese, viene utilizzata una lingua "ponte" come l'inglese, ovvero si traduce prima dal russo all'inglese e poi dall'inglese al giapponese), in quanto la lingua *pivot* crea delle situazioni di ambiguità e vi è un'alta probabilità che alcuni dettagli o informazioni vadano persi, mentre l'interlingua evita proprio che si verifichino tali circostanze (Melby, 2020:424).

La nozione di *transfer rules* (regole di trasferimento), riferita ai sistemi RBMT, apparve per la prima volta negli anni '50, prova che fin dal principio i ricercatori posero l'attenzione sulle lingue *source* e *target*: per passare da una lingua di partenza a una di arrivo, sono necessarie alcune informazioni su come tradurre gruppi di parole; è essenziale tenere conto del fatto che una frase ha una struttura complessa, quindi è preferibile, anziché considerarla interamente nel suo insieme, suddividerla in frammenti, da poter poi tradurre utilizzando regole distinte (Poibeau, 2017:22). Possono essere utilizzate delle regole di trasferimento che specificano, ad esempio, la posizione dell'aggettivo: come riporta in un esempio Poibeau (2017:22), normalmente gli aggettivi in inglese sono collocati prima del sostantivo, mentre nel francese, dopo. Per casi con strutture più complesse, in cui non vi è un'esatta corrispondenza tra la parola della lingua di partenza e la parola della lingua di arrivo, è possibile applicare delle regole più complesse (ibidem). Il trasferimento può anche essere attuato a livello semantico, affinché possa essere scelta la giusta traduzione di un determinato termine, dato un contesto di riferimento: per esempio, la parola inglese *bank*, in base al contesto, può indicare sia un istituto di credito che la riva di un fiume (ivi, 23). Il problema del dominio è evidente nella traduzione manuale, e per questo è da sempre una delle grandi sfide della traduzione automatica. Solo grazie all'osservazione di una grande quantità di dati si può arrivare a dei risultati più accurati, ottenuti, fra l'altro, con i sistemi di traduzione automatica statistici (SMT) e neurali (NMT).

I sistemi RBMT dominarono il mercato per molti anni: il primo sistema di traduzione online, gratuito, nel 1997, era infatti basato su regole (Joscelyne 1998 in Kenny, 2022:35). Tuttavia, i sistemi di questi tipo presentavano una lunga serie di problematiche, tra cui l'alto costo richiesto per il loro sviluppo, dato anche dalla presenza necessaria di linguisti altamente preparati, che dovevano occuparsi della stesura di regole specifiche per ogni coppia linguistica (Kenny, 2022:35); oltretutto, in molti casi era difficile, o addirittura impossibile, anticipare tutta la conoscenza necessaria a far funzionare i sistemi RBMT come si desiderava (ibidem). Attualmente, difatti, i sistemi basati su regole sono considerati obsoleti, e continuano a esistere solamente in sistemi ibridi o in sistemi molto vecchi. Una prova di ciò è fornita dall'Organizzazione Panamericana della Sanità (PAHO), che utilizza tuttora un sistema *rule-based*, in quanto al suo interno si cominciò a fare uso di sistemi di traduzione automatica negli anni '80, lavorando fino da allora con solamente tre lingue (Nitzke & Hansen-Schirra, 2021:23).

### **1.3.2 Corpus-based Machine Translation**

Accanto ai sistemi RBMT, esiste un altro approccio alla traduzione automatica, ovvero la *corpus-based machine translation*. La traduzione automatica *corpus-based* si occupa di esaminare una serie di tecniche, attraverso le quali, le regole che dominano il processo di traduzione possono essere ottenute automaticamente da dei corpora e applicate in un processo stocastico, al fine di ottenere delle traduzioni di testi sorgenti ancora non visti e/o analizzati (Forcada, 2010:219; Schwartz, 2018:171). È basata sui dati e dunque fa parte di quei sistemi *data-driven*. Tali sistemi si trovano alla base del pensiero della cosiddetta *machine learning*, che segna un grosso cambiamento ideologico nella ricerca sulla traduzione automatica: anziché suggerire al computer cosa fare, si lascia acquisire delle conoscenze a quest'ultimo (Kenny, 2022:36). All'interno del computer risiedono una serie di dati, di problemi, risolti nel passato, e, grazie a queste informazioni, la macchina è in grado di elaborare delle risposte ai problemi futuri, sulla base delle soluzioni passate (ibidem). Questi dati, da cui la stessa macchina "impara", consentono alla *machine learning* di distinguersi dagli altri tipi di intelligenza artificiale (ibidem). Esistono altri due tipi di sistemi basati sui dati, ovvero i sistemi SMT e NMT, che verranno analizzati in seguito (vedi §1.3.3 **Statistical Machine Translation** e §1.3.4 **Neural Machine Translation**).

Per risalire alle origini dei sistemi *corpus-based*, è necessario tornare agli anni immediatamente seguenti alla Seconda Guerra Mondiale, quando Warren Weaver cominciò a ipotizzare che la traduzione potesse essere trattata come un problema di crittografia: egli riteneva, infatti, che i nuovi metodi meccanizzati nella crittografia, considerati molto potenti, potessero ottenere risultati positivi anche quando non si sapeva che lingua fosse stata codificata (Weaver, 1947, in Schwartz, 2018:175). Quest'idea fu descritta in modo approfondito un paio di anni più tardi, quando lo stesso Weaver pubblicò un memorandum nel 1949, volto ai ricercatori interessati alla traduzione automatica, che stava nascendo e crescendo proprio a quel tempo (Schwartz, 2018:175). Weaver, a sua volta, si ispirò al lavoro di Shannon (1948), il cui approccio per un modello statistico della lingua umana, basato su un contesto limitato, poteva essere usato per disambiguare parole con molteplici possibili significati durante la traduzione automatica (ibidem).

Le regole su cui erano (e sono) basati i sistemi RBMT venivano create da linguisti, umani (Schwartz, 2018:175). Tuttavia, nel corso degli anni '80, un gruppo di ricercatori giapponesi cominciò a indagare sulla fattibilità dell'ottenimento di tali regole non più da persone, bensì automaticamente da un corpus parallelo di dati tradotti (ivi, 175-176). Nel 1984, infatti, Nagao affermò che, di fronte a una nuova frase da tradurre, i segmenti componenti la frase stessa potevano essere resi nella lingua di arrivo trovando delle traduzioni già esistenti di quegli stessi segmenti o altri simili, ovvero una traduzione basata su degli esempi (*Example-Based Machine Translation*, EBMT) (ivi, 176). La traduzione automatica basata su esempi presentava alcune somiglianze con le memorie di traduzione, sviluppate diversi anni più tardi per aiutare i traduttori umani nello svolgimento del proprio lavoro (ibidem).

Dunque, esistono due tipologie di tecnologie di traduzione automatica basate su un corpus, ovvero la *Example-Based Machine Translation* e la *Statistical Machine Translation*. Prerequisito fondamentale per entrambi gli approcci è un corpus di testi bilingui paralleli allineati alle frasi (Forcada, 2010:219). La EBMT inizialmente fu proposta da Nagao con la dicitura di “traduzione per analogia” e normalmente è composta di tre fasi (ibidem):

- *Matching*: una nuova frase, che deve essere tradotta, viene divisa in segmenti, i quali vengono confrontati con segmenti identici o simili nella lingua *source* degli esempi bilingue, presenti nel corpus (ibidem).
- *Alignment*: vengono determinati i frammenti corrispondenti nella lingua di arrivo degli esempi bilingui abbinati, per costruire una "unità di traduzione". (ibidem).
- *Recombination*: si produce, infine, una traduzione per la frase di partenza, tramite la combinazione delle precedenti unità di traduzione, nella lingua di arrivo (ibidem).

Queste tre fasi sono viste come parallele alle tre fasi di *analysis*, *transfer* e *generation* nella RBMT (Somers, 2003, in Forcada, 2010:220). Una differenza evidente tra EBMT e SMT risiede nel fatto che quando una nuova frase è identica a una già presente nel corpus di esempi, la sua traduzione, nel caso di EBMT, è uguale a quella già esistente e usata in questo modo (come avviene per le memorie di traduzione) (Forcada, 2010:220). Nei sistemi SMT, tuttavia, ciò non avviene (ibidem). I sistemi EBMT hanno un livello di pre-elaborazione del corpus di esempi bilingue, che precede il processo di traduzione vero e proprio, e che comporta l'uso di dizionari bilingue o *parser* (decodificatori) come risorse (ibidem). Vi sono anche una serie di difficoltà che sono legate alle tre fasi precedentemente analizzate: la segmentazione, infatti, deve avvenire correttamente, rappresentando essa stessa una parte cruciale per l'ottenimento di un *matching* corretto; l'allineamento non è una fase da sottovalutare, da giudicare come banale, in modo particolare quando le lingue prese in considerazione non sono tra di loro vicine; infine la ricombinazione può essere ostacolata da problematiche quali la ripetizione di parole, ecc. (ibidem).

Nel prossimo capitolo viene preso in considerazione il secondo dei due sistemi basati sul corpus, ovvero il sistema di traduzione automatica statistica.

### **1.3.3 Statistical Machine Translation**

I sistemi SMT furono considerati i più avanzati per diversi anni, in modo particolare a partire dall'inizio del 2000. La traduzione automatica statistica si sviluppò in modo indipendente dalla precedentemente nominata EBMT: fu il risultato di un fondamentale articolo accademico, pubblicato da Brown *et al.* (1988, in Forcada, 2010:220). I sistemi

di traduzione automatica statistica non sono basati su regole, bensì su dei dati raccolti: infatti, sono anch'essi dei sistemi *data-driven*.

Ci furono diversi momenti nella realizzazione della SMT. Schwartz (2018:176), infatti, nomina una prima tipologia di sistema SMT, ovvero il sistema di traduzione automatica denominato *word-based statistical machine translation*. I sistemi di traduzione automatica statistica, infatti, derivano da alcune ricerche, compiute nel corso degli anni '70 e '80 del Novecento, sul riconoscimento vocale automatico (*automatic speech recognition*, ASR) (ibidem). Le ricerche fornirono degli ottimi risultati, con una buona riuscita della correzione automatica dell'ortografia, tanto che un gruppo di ricerca dell'IBM, sotto la supervisione di Fred Jelinek, decise di applicare le stesse tecniche anche alla traduzione automatica, campo che fino a quel momento era dominato dai sistemi basati su regole RBMT (ibidem). L'IBM presentò, dunque, la prima proposta concreta di una traduzione automatica basata sulle idee teorico-informatiche di Weaver e Shannon (Brown *et al.*, 1988 in Schwartz, 2018:176); cinque anni più tardi lo stesso gruppo dell'IBM riuscì a sviluppare dei modelli sempre più sofisticati per una traduzione *word-for-word* probabilistica: infatti, la SMT si basa su un modello di probabilità, tratto da un corpus bilingue (Brown *et al.*, 1993 in Schwartz, 2018:176; Forcada, 2010:220). Questi modelli appresero le regole della traduzione parola per parola probabilistica direttamente da un corpus parallelo di testi tradotti, senza la necessità dell'intervento umano o di approfondimenti linguistici: quello proposto da IBM era un modello di traduzione automatica statistica, chiamato "Candice", che si scostava totalmente dalle regole messe a punto dalla mente umana, tipiche dei sistemi (prevalenti all'epoca) *transfer-based* e *interlingual* (Schwartz, 2018:176).

Con l'avvento dei sistemi statistici *word-based*, si crearono due principali correnti di ricerca: alcuni ricercatori proseguirono con lo studio e il miglioramento dei sistemi RBMT; altri, viceversa, continuarono a sperimentare con il nuovo approccio statistico basato su corpus, introdotto da IBM (ivi, 177). Fu proprio dall'evoluzione di questa prima tipologia di sistema SMT che, verso la fine degli anni '90, numerosi studiosi (quali Melamed 1997, Wang & Waibel 1998, Och *et al.* 1999) iniziarono a sviluppare nuovi modelli partendo dalla *word-based* SMT: l'obiettivo era quello di "catturare" traduzioni di molteplici parole adiacenti e sviluppare algoritmi per consentire la traduzione tramite questi modelli (Zens *et al.* 2002, Watanabe *et al.* 2003, Koehn *et al.*

2003, in Schwartz, 2018:177). Questo nuovo modello venne chiamato *phrase-based* SMT e, già entro la metà degli anni 2000, rimpiazzò il sistema *word-based* (Koehn *et al.* 2003, in Schwartz, 2018:177). Tuttavia, il termine *phrase-based*, ovvero “basato su frasi”, risulta fuorviante, poiché i segmenti in questione non sempre, e non necessariamente, corrispondono a ciò che si intende con “frase” nella linguistica (Kenny, 2022:37). Si preferisce, infatti, la dicitura “*n*-grammi”, con la quale si indicano sequenze di una, due, tre o un numero *n* di parole, che compaiono in modo contiguo nei dati di addestramento (si chiamano, dunque, in base al numero di parole che lo compongono, ad esempio unigramma, se si tratta di una singola parola, digrammi, se sono due, o trigrammi se tre, e così via) (*ibidem*). I sistemi *phrase-based* sono dei modelli bilingue in cui gli *n*-grammi, rilevati all’interno della parte dei dati di addestramento nella lingua di partenza, vengono sistemati all’interno di una tabella, accanto alle loro possibili traduzioni identificate nella lingua di arrivo, e ognuna di queste coppie di combinazioni linguistiche viene accompagnata da un punteggio percentuale di attinenza (*ivi*, 36). Kenny (2022:36) fornisce anche un esempio molto semplificato del funzionamento di un sistema SMT *phrase-based* (in realtà, le combinazioni possibili sono in numero nettamente maggiore, molte delle quali risultano senza senso):

	English	Probability
a me piace	I like	0.78
a me piace	I should like to	0.11
a me piace	I admire	0.11

Tabella 1: estratto di una *phrase table* che mostra le possibili traduzioni in inglese di *a me piace*, con la relativa percentuale di probabilità, tratto dal Corpus Europarl del Parlamento Europeo (Kenny, 2022:36).

Esiste anche un modello di traduzione automatica statistica di tipo monolingue: il *language model* è, infatti, un modello (o una combinazione di modelli) monolingue della lingua di arrivo (Kenny, 2022:37). Come il precedente, anche il *language model* è basato sugli *n*-grammi (*ibidem*).

I sistemi di traduzione automatica statistica prevedono la presenza di due processi distinti e separati tra loro, ovvero una fase di addestramento (*training*) e di decodifica (*decoding*) (Hearne & Way, 2011:205). Nella fase di addestramento vengono estratti sia

un modello statistico di traduzione da un corpus parallelo, che un modello statistico della lingua di arrivo da un corpus monolingue, normalmente più ampio di quello parallelo (Brown *et al.* 1990, 1993, in Hearne & Way, 2011:205). Il modello di traduzione comprende un dizionario bilingue in cui vi è una associazione per ogni possibile traduzione di una singola parola o frase del testo sorgente, con una relativa percentuale di probabilità, comprendente anche opzioni poco plausibili (Hearne & Way, 2011:205). Durante la fase di addestramento possono essere ricavate ulteriori informazioni come, ad esempio, i modelli sul riordino delle parole, la lunghezza relativa delle frasi, l'importanza relativa della traduzione vs il modello linguistico ecc. (ibidem). Questi modelli vengono successivamente usati durante la fase di decodifica, nella quale la traduzione viene trattata come un problema di ricerca (ibidem). Infatti, la fase di decodifica consiste nella ricerca, tra tutte le traduzioni possibili secondo il modello di traduzione, della soluzione più probabilmente corretta, ovvero con la percentuale di accuratezza e attinenza più alta (ivi, 205-206). I sistemi SMT puntano a generare migliaia di traduzioni ipotetiche e successivamente scegliere quella che può essere considerata come la migliore o più plausibile, lavorando, dunque, sulla base della probabilità (ivi, 206). Kenny (2022:37) tra le due fasi precedentemente elencate ne elenca un'altra, ovvero la cosiddetta fase di *tuning*, di "sintonizzazione", in cui viene assegnato un determinato peso a ogni modello, in modo da ottenere il miglior output. La fase di tuning precede quella di decodifica.

I sistemi SMT furono quelli considerati più all'avanguardia, sostituendosi quasi completamente ai precedenti sistemi RBMT, fino al 2015 (ibidem). Vi erano, tuttavia, una serie di mancanze e problematiche, la maggior parte delle quali dovuta al fatto che gli *n*-grammi relativamente corti erano usati per costruire dei modelli e che gli *n*-grammi nella stessa frase venivano tradotti come se fossero stati indipendenti l'uno dall'altro (ibidem). Inoltre, i sistemi SMT sembravano non funzionare particolarmente bene su lingue altamente flesse e agglutinanti (ibidem). Altri problemi erano da attribuirsi alla mancata traduzione di alcune parole, o all'inconsistenza, ovvero quando la stessa parola della lingua di partenza veniva tradotta in due modi diversi, a volte all'interno della stessa frase (ibidem). Gli errori prodotti dai sistemi di traduzione automatica, tuttavia, sono normalmente abbastanza facilmente prevedibili e non

possono essere migliorati o risolti fintanto che il sistema non viene allenato con dei nuovi corpora (Nitzke & Hansen-Schirra, 2021:25).

Un compito finale, che può rientrare all'interno degli obiettivi dell'addestramento di tali sistemi, è la valutazione della performance del sistema stesso: si tratta di MERT (Minimum Error Rate Training) (Hearne & Way, 2011:219). Molto spesso questa tecnica di valutazione include in essa un'altra metrica, precedentemente nominata, ovvero BLEU: si tratta di un valore calcolato automaticamente che valuta la qualità di un sistema MT, fornendo un paragone tra il testo sorgente e l'output della traduzione prodotta dal sistema (ivi, 219-220).

La tecnologia proposta dai sistemi statistici fu utilizzata da grandi aziende, tra le quali Google Translate (dal 2007), e istituzioni, come la Commissione Europea (dal 2010): anche nelle cosiddette *shared task evaluations* (valutazioni di compiti condivisi, in cui venivano messi a confronto dei sistemi, l'uno contro l'altro, per vedere quale fornisse una performance migliore in una determinata coppia linguistica e con determinati tipi di dati di addestramento), i sistemi SMT risultavano i più adatti e all'avanguardia (Kenny 2022:39). Nel 2015, però, cominciò a prendere piede un nuovo approccio, anch'esso di tipo *data-driven*, ovvero quello neurale (ivi, 38). Sebbene i sistemi di traduzione automatica statistica divennero obsoleti, contribuirono in maniera fondamentale allo sviluppo dei sistemi di traduzione automatica, dimostrando che, nonostante i vari problemi riscontrati, un sistema basato su dati, in grado di imparare da sé stesso, fosse migliore dei precedenti sistemi, basati su regole ideate dall'uomo (ibidem). Inoltre, l'enorme quantità di dati raccolti, sia monolingui che bilingui, inseriti dagli sviluppatori dei sistemi SMT, provenienti da parlamenti, organizzazioni internazionali, dal web, ecc., ha fortemente contribuito alla ricerca nel campo della traduzione automatica mondiale, consentendo i progressi che hanno reso possibile l'evoluzione di tale tecnologia (ibidem).

Nel prossimo capitolo viene analizzato nel dettaglio cosa si intende con “traduzione automatica neurale” e quali sono le sue caratteristiche specifiche.

#### **1.3.4 Neural Machine Translation**

La scalata verso il successo dell'approccio neurale iniziò attorno al 2015, surclassando il precedente sistema SMT a livello performativo. Fu proprio nel 2015, infatti, che un sistema di traduzione automatica neurale (NMT), sviluppato presso l'Università di

Stanford, si dimostrò migliore di numerosi sistemi SMT: fu questo, infatti, l'esito di un'indagine effettuata su una coppia linguistica considerata difficile, ovvero la coppia inglese-tedesco (Bentivogli *et al.* 2016 in Kenny, 2022:39). Il sistema proposto dall'Università di Stanford riscosse un enorme successo, dando inizio a quella che venne definita da Bentivogli *et al.* (2016) come “la nuova era della NMT” (Kenny, 2022:39). Nell'approccio neurale la traduzione appare, grazie alla forma in cui si presenta e alla scorrevolezza del testo, molto simile a una traduzione svolta da un traduttore umano, tanto che alcuni ricercatori inizialmente affermarono che, con la NMT, si fosse raggiunta la parità umana nella traduzione automatica (ibidem). Infatti, venne anche affermato da Marking (2016, in Kenny, 2022:39), che la traduzione neurale fosse in grado di apprendere e riprodurre metafore, espressioni idiomatiche, proponendo non una traduzione letterale, bensì riproducendo il contesto culturale adatto in un'altra lingua. Tali affermazioni possono, tuttavia, essere considerate totalmente corrette? Parzialmente lo sono, in quanto la NMT è in grado di produrre una traduzione automatica simile a quella umana, grazie all'addestramento del sistema con una grossa quantità di dati e, dunque, mediante la presenza di precedenti traduzioni di testi simili o addirittura uguali ad altri affrontati (Kenny, 2022:39). Il sistema, dunque, in realtà non sa di essere idiomatico, ma riproduce ciò che ha appreso dai dati (ibidem). Il lavoro svolto dalla traduzione neurale, quindi, è simile a quello compiuto precedentemente dalla traduzione statistica. Eppure, la NMT sembra surclassare, in qualità dell'output prodotto, la SMT, in quanto vengono usate delle rappresentazioni diverse e dei tipi di modelli<sup>2</sup> diversi nei due sistemi presi in considerazione (ibidem).

L'architettura dei sistemi NMT si differenzia da quelle precedentemente affrontate. La traduzione automatica neurale, diversamente dalla SMT, che utilizza degli *n*-grammi, si ispira al cervello umano e considera la traduzione come un'attività che coinvolge numerose operazioni, svolte da sistemi matematici chiamati “reti neurali artificiali”, i quali, presa una determinata frase, la trasformano in una serie di numeri (Pérez-Ortiz *et al.*, 2022:141). Queste reti sono formate da migliaia di unità individuali, chiamate neuroni artificiali, e svolgono dei compiti computazionali specifici, volti a risolvere un

---

<sup>2</sup> Un modello informatico è una rappresentazione matematica astratta di un evento, di un sistema o di un fenomeno della vita reale (Kenny, 2022:40). Un modello ha diversi usi, uno dei quali è quello di prevedere la risposta a un problema nuovo, quindi non affrontato precedentemente: ad esempio, un modello computazionale di traduzione dovrebbe essere in grado di prevedere una frase nella lingua di arrivo, data una frase nella lingua di partenza non analizzata in precedenza (ibidem).

problema, fornendo una stima delle probabilità esistenti (Pérez-Ortiz *et al.*, 2022:145-6; van Genabith, 2020:59). Tali neuroni, per funzionare e svolgere determinati compiti, vengono attivati in base agli stimoli che ricevono dagli altri neuroni e dal peso delle connessioni esistenti tra i neuroni stessi (Pérez-Ortiz *et al.*, 2022:144-146). La rete neurale, su cui sono basati suddetti sistemi, è di tipo *end-to-end*, ovvero continua, e rileva testi bilingue allineati; questa rete, data una frase input da tradurre, viene normalmente addestrata per massimizzare la probabilità di ottenere una frase di arrivo senza l'aggiunta di ulteriori informazioni linguistiche esterne (Castilho *et al.*, 2017:110). Nella costruzione di un sistema NMT è fondamentale, innanzitutto, determinare la sua architettura, ovvero da quanti neuroni è formata, come sono tra di loro connessi, quali ricevono degli input esterni e quali invece si occupano della produzione dell'output (Pérez-Ortiz *et al.*, 2022:146). Tuttavia, la reale computazione è svolta in base ai pesi di tutte le connessioni nella rete (*ibidem*).

Castilho *et al.* (2017:110) cita anche le reti neurali profonde, *deep neural networks* (DNN): queste reti neurali profonde vengono applicate per la costruzione di modelli di codificazione-decodificazione *end-to-end* (Kalchbrenner & Blunsom, 2013; Sutskever *et al.*, 2014; Cho *et al.*, 2014, in Castilho *et al.*, 2017:110). Esistono, dunque, due fasi nei sistemi NMT: codifica (*encoding*) e decodifica (*decoding*). Durante la fase di codifica il sistema, per tradurre la frase, assegna una rappresentazione neurale, o *embedding* (incorporamento), a ognuna delle parole del testo sorgente, in modo isolato (Pérez-Ortiz, 2022:143). Tali rappresentazioni neurali vengono unite e combinate per produrre, a livello di frase, una rappresentazione simile (*ibidem*). Quando vengono unite e combinate, subiscono un adattamento di significato in base al contesto, in quanto viene interpretato sia il significato letterale che quello idiomatico (*ibidem*). Nella fase di decodifica, invece, le rappresentazioni sono a livello di frase e predicano, una a una, le parole nella frase della lingua di arrivo (*ibidem*). Esiste, alla base di questa due fasi, un meccanismo di attenzione (introdotto per la prima volta da Bahdanau *et al.*, 2015, come riportato da Castilho *et al.*, 2017:110), che condiziona la previsione di ogni parola in uscita a una distribuzione delle parole in ingresso (lavora nello stesso modo, a grosso modo, della funzione di allineamento) (Koehn *et al.*, 2020:119). La decodifica, dunque, procede parola per parola, e, a ogni step nel processo di decodificazione, viene calcolata una percentuale di probabilità di distribuzione sulle possibili parole seguenti,

scegliendo, successivamente, la parola con il punteggio più alto, essendo con molta probabilità la più corretta, come avveniva anche nei sistemi SMT (ibidem). Con decodifica si intende anche in questo caso, infatti, la produzione di un output nella lingua di arrivo.

In un modello neurale l'algoritmo è esposto a enormi quantità di dati paralleli, con cui addestrare il sistema. Durante l'addestramento viene determinato il peso delle connessioni tra i neuroni costituenti della rete neurale, affinché si possa ottenere un output il più possibile vicino al significato originale, tenendo conto dei dati di addestramento (Pérez-Ortiz, 2022:148). L'addestramento inizia con dei pesi casuali o presi da una rete neurale in grado di risolvere un compito simile (ibidem). Durante l'addestramento poi, i pesi subiscono delle modifiche, affinché il valore di una funzione di errore (o funzione di perdita), che misura quanto gli output effettivi si discostano dalla soluzione desiderata, sia il più basso possibile, ovvero che la traduzione risulti il più precisa possibile (ibidem). Gli algoritmi di addestramento sono chiamati anche algoritmi di apprendimento e calcolano piccole correzioni o aggiornamenti da apportare ai pesi, fino a quando non si ottiene una determinata performance, o fino a quando la funzione di errore risulta minima o sufficientemente bassa (ivi, 148-149). Questo tipo di addestramento è chiamato *gradient descent* (discesa del gradiente) ed è volto a trovare i pesi migliori (o dei buoni candidati) (ivi, 149). Il tasso di apprendimento è un parametro che regola queste variazioni dei pesi; solitamente è più alto nei primi step di apprendimento, e si riduce durante l'addestramento, fino ad arrivare al punto in cui il peso è determinato, e in quel momento l'addestramento arriva al termine (ibidem).

Un altro processo cognitivo fondamentale è la generalizzazione, presente, innanzitutto, all'interno della mente umana e animale, e successivamente applicata ai sistemi di traduzione neurale (ibidem). Si tratta del processo cognitivo grazie al quale si è in grado di rintracciare ciò che si è imparato nel passato e riutilizzare questi dati per delle situazioni non identiche a quelle già affrontate, ma che presentano delle similarità con esse (ibidem). La generalizzazione sta alla base della NMT, in quanto, se allenato con un numero sufficiente di dati, un sistema di traduzione neurale è in grado di trovare delle soluzioni sulla base di situazioni simili, incontrate precedentemente in testi monolingui o bilingui, proponendo delle traduzioni corrette.

Nella NMT viene usato un particolare tipo di rappresentazione, il “vettore”. Il vettore è formato da una lista di numeri di dimensioni fisse (Kenny, 2022:41). Ad esempio, come suggerito da Kenny (2022:41), la parola “mela” potrebbe essere rappresentata da un vettore come [1.20, 2.80, 6.10] (ibidem). Il vettore rappresenta una relazione tra le parole: ciò vuol dire che i vettori sono simili tra parole con significati tra loro analoghi o che ricorrono in contesti affini (ad esempio, la parola “pera”, sarà rappresentata da un vettore simile a quello della parola “mela”, in quanto sono entrambi dei frutti) (ivi, 42). Queste rappresentazioni basate su vettori vengono anche denominate *word embeddings* (incorporamenti di parole) (Mikolov *et al.*, 2013, in Pérez-Ortiz, 2022:150). La quantità di numeri che rappresentano una parola dipende dal numero di dimensioni preso in considerazione: i sistemi NMT hanno degli incorporamenti, con centinaia di dimensioni (ivi, 151). Gli incorporamenti di parole sono costruiti a strati: una rete neurale artificiale costituita da una molteplicità di strati è una rete neurale profonda, già nominata in precedenza (ivi, 147).

Numerose ricerche hanno provato che i sistemi neurali sono migliori sotto diversi aspetti, riuscendo a svolgere con successo un elevato numero di compiti; oltretutto, in seguito a valutazioni umane su piccola scala e valutazioni automatiche, è stato dimostrato un aumento della qualità nei sistemi NMT, rispetto a quelli statistici (Bahdanau *et al.*, 2015; Jean *et al.*, 2015; Bentivogli *et al.*, 2016; Wu *et al.*, 2016, in Castilho *et al.*, 2017:110). I sistemi di traduzione automatica neurale, fin da subito dopo il loro esordio, si dimostrarono superiori ai sistemi di traduzione automatica statistica, dando degli ottimi risultati, come dimostrato da alcuni studi di Wu *et al.* (2016), con determinate lingue quali inglese, spagnolo, francese, cinese semplificato (Castilho *et al.*, 2017:110). Bentivogli *et al.* (2016), poi, riscontrarono una diminuzione nella necessità di effettuare il post-editing nel caso di output prodotti da sistemi neurali, rispetto ai sistemi SMT (ibidem). I sistemi neurali si distinguono, in quanto non presentano molti errori di flessione e di ordine degli elementi all’interno della frase in molte coppie linguistiche, ma spesso vi sono delle omissioni di parole o degli altri errori che facilmente riescono a scappare all’occhio umano, in quanto il testo risulta fluente (ibidem). Nonostante l’aumento della qualità, la NMT continua, dunque, a nascondere delle insidie e a produrre un output non privo di errori, soprattutto se il sistema non è stato addestrato con un numero sufficiente di dati, o se prendono in considerazione delle

combinazioni di lingue parlate solamente da una minoranza di popolazione, o considerate più “rare”, di cui non si possiedono grandi quantità di dati: in questo caso vi è un alto rischio che l’output fornito sia di bassa qualità. Gli errori prodotti da questi sistemi vengono affrontati ed elencati successivamente al paragrafo **§2.4.1 Errori tipici dei sistemi di traduzione automatica neurale**.

#### **1.4 Traduzione Assistita**

Accanto alla *machine translation*, vi è un’altra tipologia di traduzione, che si serve anch’essa di un sistema, ma è volta a fornire determinati strumenti al traduttore, utili durante il suo lavoro. Si tratta della *computer-aided translation*, ovvero la traduzione assistita, che, come suggerisce la sua denominazione, aiuta il lavoro umano, senza sostituirlo completamente. Vengono di seguito analizzate le principali caratteristiche della traduzione assistita, il suo funzionamento e i diversi strumenti.

##### **1.4.1 Computer-Aided Translation**

Con il termine *Computer-Aided Translation* (CAT) si intende un software volto ad aiutare e assistere un traduttore (umano) nel processo di traduzione (Bowker & Fisher, 2010:60). Differisce dalla *machine translation*, in quanto il risultato finale della traduzione è da attribuirsi primariamente al traduttore, che però viene aiutato dal software durante il processo di traduzione. Lo scopo, dunque, non è quello di sostituire la traduzione umana, ma, al contrario, aiutarla (O’Hogan, 2009:48). La nascita dei sistemi CAT è frutto di una conseguenza naturale in un mondo altamente globalizzato: tali sistemi, infatti, vanno incontro all’esigenza di accorciare i tempi di traduzione e i costi derivanti da essa; inoltre, la quantità di contenuti (di partenza) digitalizzati aumenta di giorno in giorno, e vi è, dunque, il bisogno di tradurre sempre più testi, ma nel minor tempo possibile (ibidem).

La nascita dei sistemi di traduzione assistita risale agli anni ’60 del Novecento, quando i ricercatori cominciarono a creare delle banche dati di termini, immagazzinate all’interno di un computer (Bowker & Fisher, 2010:60). Fu solo tra gli anni ’70 e ’80 del secolo scorso che si cominciò a parlare di *CAT tool*, quindi degli strumenti CAT moderni, che non si occupavano solamente di immagazzinare banche dati terminologiche e informazioni all’interno di computer, ma anche della ricerca e del recupero di queste (ibidem). Infatti, negli anni ’80 Martin Kay evidenziò la necessità, a suo dire, di riesaminare, all’interno del processo di traduzione, il ruolo rispettivamente

di uomini e macchine, proponendo di riportare l'attenzione verso gli uomini, mentre in quel momento storico ci si stava focalizzando su una traduzione automatizzata, svolta dalle macchine (Schwartz, 2018:182). In modo particolare, proponeva una cooperazione tra le due parti, esattamente quello di cui si occupano gli strumenti CAT (ibidem). Martin Kay propose quello che definì come un "amanuense del traduttore", ovvero ciò che oggi conosciamo con il termine di *Translator Environment Tool* (TEnT) (Bowker, 2020:456). Il TEnT può essere descritto come una sorta di stazione di lavoro per il traduttore o una suite di strumenti, in cui quest'ultimi sono resi disponibili al traduttore per aiutarlo durante le varie fasi del suo lavoro e sotto diversi aspetti (ibidem). Il sistema di memorie di traduzione, *Translation Memory system* (TM), rappresenta lo strumento chiave di questo ambiente (TEnT) (O'Hagan, 2009, in Bowker, 2020:456). Le memorie di traduzione funzionano, molto spesso, in stretta associazione con un *Terminology Management System*, un sistema di gestione terminologica (Bowker & Fisher, 2010:60). I *Terminology Management System* sono degli strumenti usati per immagazzinare informazioni terminologiche e recuperarle da un *termbase*, ovvero un database contenente i singoli termini (ivi, 61). Gli strumenti CAT comprendono qualsiasi strumento computerizzato che possa aiutare il traduttore nello svolgimento del proprio lavoro (Bowker, 2002:6, in O'Hogan, 2009:48; Bowker, 2009:289). Essi, dunque, variano molto, spaziando da applicazioni di uso generale, come elaboratori di testi, motori di ricerca su Internet, software di riconoscimento ottico dei caratteri (*Optical Character Recognition software* – OCR), a strumenti più orientati alla traduzione, quali quelli dedicati all'analisi del corpus, sistemi di estrazione terminologica e i precedentemente nominati sistemi di gestione terminologica (*Terminology Management System*), o ancora dizionari elettronici multilingue (O'Hagan, 2009:48).

Si può affermare con sicurezza che oggigiorno la maggior parte dei traduttori si serve dei sistemi CAT nello svolgimento del proprio lavoro. Questa tecnologia è divenuta, fin dalla sua introduzione, la predominante nella traduzione scientifica e tecnica e nella localizzazione, nella quale viene utilizzata al fine di aumentare la produttività e un'efficacia nei costi, oltre a tentare di migliorare la qualità della traduzione stessa (O'Hogan, 2009:48). Gli strumenti CAT sono stati fin dai subito adottati da agenzie di traduzione, organizzazioni internazionali, enti, ecc. (Bowker & Fisher, 2010:62).

Ciononostante, i sistemi CAT non funzionano sempre correttamente: come verrà ribadito nella prossima sezione, dedicata alle memorie di traduzione, questi tipi di sistema funzionano bene su testi ripetitivi o simili, nel contenuto e nella struttura, ad altri precedentemente tradotti; per quanto riguarda testi difficilmente “prevedibili”, ovvero, ad esempio, testi di tipo letterario, o con un contenuto particolarmente creativo, questi non sono adatti al tipo di tecnologia proposto dagli strumenti CAT (ivi, 62-63). La qualità della traduzione ottenuta grazie all’utilizzo di strumenti CAT dipende comunque dal traduttore umano, che ha il compito di controllare l’esattezza dei suggerimenti proposti dallo strumento (ivi, 63).

Un altro argomento dibattuto è come questa tecnologia influisca sulla figura del traduttore e sul suo status, la sua remunerazione e i suoi diritti di proprietà intellettuale (ibidem). Inizialmente, molti clienti attribuivano meno valore al lavoro di un traduttore che utilizzava degli strumenti CAT: essi sostenevano di poter pagare meno un traduttore che utilizzasse tali tecnologie, in quanto molto lavoro veniva “suggerito” dal sistema stesso, velocizzando le tempistiche (ibidem). I clienti, inoltre, potrebbero dimostrarsi ancora più esigenti nel caso in cui le TM di cui ci si deve servire per il lavoro vengono procurate dal cliente stesso, il quale, a volte, fornisce pure un testo pre-tradotto (ibidem). In ogni caso, anche se si ottengono dei suggerimenti esatti, ovvero degli *exact match*, il contenuto necessita comunque di essere controllato e sistemato, per assicurarsi che non ci siano errori (ibidem). Dunque, i tradizionali metodi di compensazione, intesi come il pagamento per parola o per pagina, cominciarono a non essere considerati più adeguati, dibattito che, al giorno d’oggi, continua a presentarsi anche nel post-editing (ibidem).

Vi sono, infine, delle discussioni dal punto di vista etico, finanziario e legale attorno al possesso e alla condivisione dei dati processati e salvati all’interno di questi strumenti: per esempio, il possesso dei contenuti delle TM, riguardanti il testo sorgente e la sua traduzione, dovrebbe presumibilmente appartenere al cliente (ibidem). Tuttavia, ciò è in dubbio quando un traduttore voglia fare libero uso dei dati, registrati all’interno di un database.

#### **1.4.2 Translation Memory Systems**

Le *Translation Memories* (TM), ovvero le memorie di traduzione, create appositamente per aiutare il traduttore umano, sono divenute uno strumento indispensabile per

numerosi traduttori. L'idea venne negli anni '90 proprio ad alcuni traduttori, attivi nel settore della localizzazione software, trovatisi a tradurre testi estremamente ripetitivi o che ripetessero parola per parola intere sezioni di versioni precedenti di un documento: si trattava, in modo specifico, di manuali, libretti di istruzioni, ecc. (Kenny, 2022:30). Ad esempio, i traduttori si trovavano a dover tradurre interamente e nuovamente manuali di software, che dovevano essere aggiornati a ogni nuovo rilascio del software stesso (ibidem). Una soluzione all'evidente monotonia di tale lavoro, che portava con sé, oltretutto, una perdita di tempo prezioso per il traduttore, fu proprio l'invenzione di questi strumenti (ibidem).

Una *translation memory* è, dunque, lo strumento che permette ai suoi utilizzatori di immagazzinare dei testi precedentemente tradotti, consultandoli facilmente per un potenziale riutilizzo (Bowker & Fisher, 2010:61). I testi *source* e *target*, in formato di *bitext*, devono essere immagazzinati nel database di una TM: il testo sorgente viene diviso quindi in segmenti (solitamente corrispondenti alle singole frasi o altri elementi come titoli/intestazioni o i contenuti delle celle all'interno di tabelle), ognuno dei quali è collegato al corrispondente segmento, rappresentante la sua traduzione (Bowker & Fisher, 2010:61; Kenny, 2022:30). La traduzione proposta, nella parte dedicata alla lingua di arrivo, è frutto di un *match*, una corrispondenza, esistente tra il segmento nella lingua di partenza e le memorie di traduzione salvate. Si creano, in questo modo, diversi tipi di *match* tra i segmenti del testo sorgente e del testo di arrivo: *exact match*, *full match*, *fuzzy match*, *sub-segment match*, *term match*, *no match* (Bowker & Fisher, 2010:61). Se si ottiene una corrispondenza esatto (*exact match*), o una molto simile, il segmento corrispondente nella lingua di arrivo può essere riutilizzato dal traduttore, in alcuni casi senza la necessità di modifica (Kenny, 2022:30). Di seguito viene riportata una tabella riassuntiva dei diversi tipi di *match* esistenti e le loro caratteristiche (da Bowker & Fisher, 2010:61):

<i>Exact match</i>	Si ottiene quando un segmento, proveniente da un nuovo testo, è totalmente identico, in ogni sua parte, a uno già presente nel database della TM.
<i>Full match</i>	Un segmento è identico a uno già presente nel database della TM, salvo per date, nomi propri, ecc.
<i>Fuzzy match</i>	Un segmento del nuovo testo presenta un grado di somiglianza

	con un segmento presente nel database della TM. A questo tipo di <i>match</i> viene assegnato un valore percentuale, dall'1% al 99%. Normalmente più alta è la percentuale proposta dal software, più utile e corretto è quel suggerimento. Il traduttore può stabilire la soglia entro la quale un <i>match</i> è considerato utile, ma diversi sistemi propongono di default una soglia tra il 60% e il 70%.
<i>Sub-segment match</i>	Un pezzo di testo contiguo all'interno di un segmento del nuovo testo è identico a un pezzo memorizzato nel database TM.
<i>Term match</i>	Un termine trovato nel nuovo testo corrisponde a una voce del <i>termbase</i> nel <i>terminology management system</i> integrato del sistema TM.
<i>No match</i>	Nessuna parte di un segmento, proveniente da un nuovo testo, corrisponde al database della TM o al <i>termbase</i> . Il traduttore deve dunque tradurre da zero; tuttavia, la nuova traduzione prodotta può essere inserita all'interno della TM, per un utilizzo futuro. Si ha un <i>no match</i> anche nel caso in cui, all'interno del sistema CAT utilizzato, sia stato impostato di considerare come un <i>no match</i> qualsiasi <i>match</i> sotto una certa percentuale (ad esempio, se viene impostato un <i>match</i> minimo del 30%, vengono considerati come <i>no match</i> tutti quelli che presentano una percentuale inferiore, dal 29% in giù).

Tabella 2: tipologie di match possibili, normalmente visualizzati nelle TM (Bowker & Fisher, 2010:61).

Nel corso del tempo le TM sono cresciute in maniera esponenziale: infatti, alcune aziende, che fin dalla nascita di tale tecnologia hanno previsto l'uso di sistemi CAT, possiedono delle memorie di traduzione contenenti centinaia di migliaia di *translation unit* (TU), unità di traduzione, ovvero dei segmenti nella lingua di partenza, allineati ai rispettivi segmenti nella lingua di arrivo (Kenny, 2022:30). Ogni TU è presentata in un formato particolare: si tratta di file XML, il cui formato è denominato *translation memory exchange*, meglio conosciuto con il suo acronimo "tmx" (ivi, 31). Di seguito viene riportato un esempio di una *translation unit* basata su un titolo (sia in inglese che

in tedesco), tratta da una TM consistente in dati provenienti dal sito web del Parlamento Europeo (Kenny, 2022:31):

```
<tu>
<tuv xml:lang="EN">
<seg>A common blacklist for unsafe airlines</seg>
</tuv>
<tuv xml:lang="DE">
<seg>Unsichere Luftfahrtunternehmen kommen auf eine schwarze
Liste</seg>
</tuv>
</tu>
```

I codici racchiusi tra il simbolo minore e maggiore, come ad esempio `<tu>` e `</tu>`, sono chiamati *tag* e nella fattispecie servono a indicare l'apertura (inizio) e la chiusura (fine) di ogni unità di traduzione; i *tag* `<tuv>` e `</tuv>` indicano, invece, l'apertura e la chiusura di ogni variante presente all'interno della TU, intese come le due varianti linguistiche inglese (EN) e tedesca (DE); infine, i *tag* `<seg>` e `</seg>` aprono e chiudono il segmento in una determinata lingua (ibidem).

Le memorie di traduzione si rivelano utili anche sotto altri aspetti: sono infatti in grado di permettere alle agenzie di traduzione di controllare (e in alcuni casi ridurre) i costi della traduzione stessa e di aumentare la competitività dell'azienda nel mercato (ibidem). Alcune organizzazioni internazionali, una tra le quali l'Unione Europea, da anni si servono di tale tecnologia, creando delle TM molto vaste e comprendenti una varietà di lingue diverse, oltretutto rendendo pubblicamente possibile la consultazione e, dunque, il loro eventuale riutilizzo da parte di altri traduttori, che possono integrarle alle proprie TM (ibidem). Infatti, una maniera per arricchire le memorie di traduzione è proprio servirsi di dati già presenti, di dominio pubblico.

Le memorie di traduzione possono essere considerate come un particolare tipo di corpus parallelo, e, nei casi in cui la traduzione sia creata senza l'uso di una TM, i testi tradotti possono essere allineati con il corrispettivo testo sorgente dopo la traduzione (ibidem). In un esempio fornito da Koehn (2005, in Kenny, 2022:31-32), il ricercatore menziona che il corpus multilingue del Parlamento Europeo, chiamato "Europarl Corpus", fu creato allineando, l'una con l'altra, le versioni dei procedimenti (*proceedings*) nelle varie lingue, ricavate dal web. I corpora paralleli allineati non si presentano in formato tmx, ma spesso sono sottoforma di file molto lunghi, composti da

migliaia o addirittura milioni di righe, ognuna delle quali è occupata da una singola frase, la cui posizione all'interno del file corrisponde alla sua traduzione in un altro file (quindi alla riga  $x$  del file nella lingua di arrivo è contenuta la traduzione della riga  $x$  nel file della lingua di partenza) (ivi, 32).

Le memorie di traduzione non servono solamente a fornire un aiuto al traduttore umano durante l'utilizzo degli strumenti CAT, ma possono anche essere integrate all'interno dei sistemi di traduzione automatica (vedi **§2.5 Post-editing nei translation memory systems**). Le TM sicuramente offrono un vantaggio in termini di tempo, perché permettendo di ritrovare con facilità non solo ciò che è stato in precedenza tradotto, ma anche proporre delle soluzioni di traduzione per segmenti non totalmente uguali ai precedenti; consentono di evitare la traduzione di porzioni di testo ridondanti; è raro che non vi sia la traduzione di un intero segmento o anche di una sola parola facente parte del segmento, in quanto il sistema fornisce sempre il tipo di *match* ottenuto e il traduttore può ricontrollare il suggerimento proposto dal sistema; inoltre, è difficile che i file vengano danneggiati, perché il sistema estrae già i testi che devono essere tradotti (Nitzke & Hansen-Schirra, 2021:44). Tuttavia, lavorare con memorie di traduzione comporta una serie di rischi e svantaggi: non è indicato utilizzare memorie di traduzione per qualsiasi tipo di testo, perché, come visto in precedenza, le TM si adattano meglio a testi come manuali ecc., ovvero testi tecnici, abbastanza ripetitivi nel contenuto, dunque, non sono adatti a testi particolarmente creativi; oltretutto, normalmente è possibile leggere il testo solamente segmento per segmento, e, nel caso in cui si ottengano diversi *exact* e *fuzzy match*, potrebbe risultare difficile la comprensione del contenuto della frase singola e piccoli (o grandi) errori potrebbero sfuggire (ivi, 44-45). Ciononostante, di norma i vantaggi superano gli svantaggi e i traduttori spesso ripongono maggiore fiducia in questa tecnologia che non nella traduzione automatica, giudicando le memorie di traduzione più "controllate" rispetto alla MT.



## II. Post-editing

### 2.1 Cosa si intende con “post-editing”

Con il termine “post-editing” (PE o MTPE, ovvero Machine Translation Post-Editing), ci si riferisce all’attività di revisione di un testo, precedentemente tradotto meccanicamente da un sistema di traduzione automatica (O’Brien, 2021:177). Il traduttore specializzato e formato nel campo del post-editing è chiamato “post-editor”. Egli si occupa di confrontare il testo originale con la traduzione “grezza” (*raw output*), prodotta dal sistema di traduzione automatica (ibidem). Il post-editing differisce dalla normale traduzione manuale in quanto il post-editor si dedica alla correzione di eventuali errori (quali omissioni, aggiunte non richieste, errori grammaticali e stilistici, ecc.) presenti nell’output prodotto dalla MT (ibidem). Come segnalato da O’Brien (ibidem), è dagli anni ‘80 del secolo scorso che si comincia a parlare di post-editing, mettendolo effettivamente in pratica. In realtà, la prima menzione al post-editing nella letteratura risale agli anni ‘50 del ‘900, ovvero attorno allo stesso periodo in cui fu avanzata la prima proposta di utilizzare sistemi per la traduzione automatica (in questo periodo chiamata anche *mechanical translation*) (Garcia, 2012:294, in Koponen *et al.*, 2021:2). I primi a adottare questa nuova tecnologia di traduzione furono grandi organizzazioni quali la Commissione Europea e l’Organizzazione Sanitaria Panamericana (Pan-American Health Organization) (Koponen *et al.*, 2021:2). Con lo sviluppo di nuove tipologie di traduzione automatica, sempre più accurate e precise, anche il post-editing si radicò e si diffuse ampiamente. Nel 2017 fu infatti pubblicato uno standard internazionale a esso dedicato: lo standard ISO 18587 (ibidem).

Do Carmo e Moorkens (2021:36) sottolineano come, per indicare questa attività, è da preferirsi il termine “post-editing”, per l’appunto, a un altro termine esistente, “post-edition”, in quanto l’uso del gerundio permette di evidenziare che si tratta di un processo: parlando di post-editing risulta infatti fondamentale concentrarsi sulle azioni che sono parte del processo di post-editing stesso. I due autori forniscono delle ulteriori indicazioni sulle differenze fondamentali tra quello che si intende con il termine “traduzione<sup>3</sup>” e quello che si intende con il termine “post-editing”. Il lavoro del traduttore si basa sulla lettura del testo sorgente e sulla traduzione del testo stesso; il

---

<sup>3</sup> Se non diversamente specificato nel testo, con il termine “traduzione” si farà riferimento alla traduzione intesa come la traduzione manuale, effettuata da un traduttore professionista umano, opposta al post-editing.

lavoro, quindi, di per sé risulta fluido, scorrevole, e si interrompe solamente nel caso in cui si presenta un problema: si tratta, dunque, di un processo di creazione (ibidem).

Per quanto banale possa apparentemente sembrare, il lavoro di post-editing è tutt'altro che semplice: non essendo prodotta da zero dal post-editor stesso, infatti, il PE della traduzione richiede una maggiore attenzione ai dettagli, anche ai più minimi, e un controllo accurato del lessico e dell'aderenza al testo originale.

Risulta fondamentale sottolineare che non tutti i testi sono adatti a essere tradotti tramite l'uso di un sistema di traduzione automatico: per la traduzione di manuali, di istruzioni per l'uso e di contratti, la traduzione automatica si addice particolarmente (Arnhold *et al.*, 2017: 221-222)<sup>4</sup>. Più un testo è creativo e stilisticamente complicato, meno sarà adatto alla traduzione automatica, che funziona in modo migliore per testi ridondanti e con un linguaggio controllato (ivi, 222). La scelta di effettuare il post-editing, dunque, non deve essere sopravvalutata, ma, al contrario, accuratamente ponderata.

### **2.1.1 Post-editing e revisione a confronto**

Sebbene il post-editing possa apparentemente sembrare equiparabile alla revisione, queste due pratiche non sono da considerarsi come equivalenti, ed è fondamentale identificare le loro particolarità. Errori lessicali, sintattici e grammaticali sono spesso presenti all'interno di traduzioni effettuate attraverso l'uso di un sistema di traduzione automatica, mentre è molto raro che errori di questo genere si trovino all'interno di traduzioni effettuate dall'uomo (Nitzke & Hansen-Schirra, 2021:10).

Come precedente sottolineato, il post-editing ha cominciato a essere realmente messo in pratica solamente verso la fine del '900. La revisione della traduzione, al contrario, ha, in Europa, origini antiche: probabilmente il primo esempio di revisione risale al IV secolo, quando Papa Damaso I commissionò a San Girolamo di revisionare la versione all'epoca esistente della Bibbia (Koponen *et al.*, 2021:1). Il vero successo dell'attività di revisione e la sua effettiva applicazione nell'industria dell'editoria avvennero, tuttavia, nel corso del secolo scorso, quando, a partire dalla fine degli anni '70, furono pubblicati diversi manuali e libri di testo contenenti indicazioni pratiche e pedagogiche, oltre a suggerimenti e studi specifici nel campo della revisione (ibidem). Il lavoro di revisione

---

<sup>4</sup> Vedi § 2.4.1 **Errori tipici dei sistemi di traduzione automatica neurale** per un approfondimento sulle diverse tipologie di sistemi di traduzione più moderni, attualmente utilizzati.

subì un maggiore riconoscimento e istituzionalizzazione da parte dell'International Organization for Standardization (Organizzazione Internazionale per la Standardizzazione - ISO) con la pubblicazione, rispettivamente nel 2006 e nel 2015, di due standard: ISO 15038 e ISO 17100 (ibidem). Nel loro contenuto si afferma che la revisione è un'attività obbligatoria nell'ambito della traduzione: deve essere effettuata da una persona esterna, diversa dal traduttore del testo stesso e consiste in un'esaminazione bilingue del contenuto della lingua d'arrivo, confrontato con quello della lingua del testo di partenza, assicurandosi che sia adeguato al fine proposto (International Organization for Standardization, 2015:2, in Koponen *et al.*, 2021:1).

La revisione, quindi, è una forma di controllo del lavoro di traduzione sia di altri traduttori che, in realtà, anche di sé stessi (*self-revision*), con il fine di individuare e correggere errori di diverse tipologie e suggerire una versione alternativa della traduzione, eseguendo un attento controllo del testo sorgente e del testo di arrivo (Koponen *et al.*, 2021:1-2).

All'interno della revisione, Mossop (2007:125, in Koponen, 2016b:16) individua le principali categorie di errori da controllare, che possono essere suddivisi in quattro gruppi: errori di trasferimento di significato, di contenuto, di lingua e di presentazione (ad esempio, il layout). Queste tipologie di errori dovrebbero essere controllate anche quando ci si occupa di MTPE (ad esempio, la correttezza delle frasi dal punto di vista grammaticale, di significato o idiomatico). Tuttavia, sussistono delle differenze: in base alle affermazioni di Koby (2001:6-7, in Koponen, 2016b:16), una delle maggiori differenze tra PE e revisione risiede nel fatto che il traduttore (umano) e il revisore possiedono entrambi un background condiviso di conoscenze extratestuali di culture, di testi e di lingue a cui poter attingere, ma queste conoscenze non appartengono ai sistemi di traduzione automatica. Il revisore, principalmente, verifica omissioni, fraintendimenti involontari compiuti dal traduttore umano, i quali potrebbero derivare da una comprensione incorretta del testo (che porta a una conseguente difficoltà di comprensione del testo da parte del pubblico); i sistemi di traduzione automatica, invece, commettono più errori localizzati relativi a parole e strutture tradotte in modo scorretto (ibidem).

Ciononostante, revisione e post-editing possiedono delle somiglianze. Sia il traduttore/revisore che il post-editor, infatti, si trovano davanti a due possibili opzioni

nei loro rispettivi lavori: possono riscrivere da zero, completamente, un passaggio di testo scorretto, oppure revisionarlo e applicare delle modifiche. La seconda opzione, tuttavia, è da considerarsi, in alcuni casi, la scelta più complessa, in quanto richiede, innanzitutto, l'individuazione del problema, e, in seguito, un metodo efficace e soddisfacente di correzione di esso (Englund Dimitrova, 2005:31, in Koponen, 2016b:16). Secondo alcuni studi compiuti all'inizio del XXI secolo da Koby (2001:10, in Koponen, 2016b:16), il lavoro del post-editor sarebbe più semplice rispetto a quello del revisore, in quanto, davanti alla scelta se riscrivere completamente un segmento o rivederlo, i post-editor sostengono di avere meno inibizioni rispetto ai revisori, affermando di riuscire ad apportare cambiamenti importanti in modo più rapido ed efficace. Su queste affermazioni si potrebbe discutere a lungo: molti traduttori sono tutt'oggi ancora ostili al dedicarsi al post-editing perché ritenuta una perdita di tempo.

Il PE è, dunque, da considerarsi più come una forma specializzata di traduzione che di revisione: se fosse una forma di revisione, l'output prodotto dalla MT dovrebbe essere praticamente perfetto, presentando una traduzione completa. Ciò non risulta possibile in quanto la MT non produce una traduzione finita, ma semplicemente un output, che si può considerare come una serie di suggerimenti o ipotesi per la traduzione del testo: la responsabilità del risultato finale della traduzione risiede, infatti, nelle mani del post-editor (Do Carmo & Moorkens, 2021:40). Ciononostante, alcuni autori si schierano a favore della teoria che equipara il post-editing alla revisione. Tesi a favore di questa teoria sono supportate da autori quali Koehn (2009) e Ortiz-Martínez et al. (2016) (citati in Do Carmo & Moorkens, 2021:40), che sostengono che il post-editing sia un lavoro più di lettura che di scrittura, come lo è la revisione, poiché diversi studi dimostrano che il post-editor spende più tempo in pause, ovvero nella lettura di un testo, che sulla tastiera, a scrivere o correggere (quindi come avviene per i revisori) (ibidem).

Sono in numero nettamente maggiore, tuttavia, le tesi a sostegno della differenziazione tra due professioni e la loro identificazione come due lavori separati. I traduttori stessi si rifiutano di comparare revisione e post-editing, affermando che quest'ultimo richiede un maggiore sforzo, oltre a una quantità di traduzione maggiore rispetto a quella richiesta dalla revisione, implicando più scrittura che lettura (Do Carmo & Moorkens, 2021:40). È importante sottolineare che alla fine del processo di post-editing non necessariamente si ottiene un testo revisionato: a seconda del livello di post-

editing richiesto dal cliente si possono ottenere risultati diversi, più o meno accurati, e nella maggioranza dei casi il post-editor deve spendere solo il tempo necessario sul testo, evitando eccessive modifiche, se non esplicitamente richiesto dal cliente (ibidem).

Krings (2001, citato da Do Carmo e Moorkens, 2021:42), nei suoi studi, sostiene che il PE richiede una maggiore attenzione nella lettura e una scrittura più complessa rispetto alla revisione e alla traduzione: la lettura, in effetti, subisce un ulteriore rallentamento se il contenuto dell'output della MT presenta una qualità di livello medio, necessitando, dunque, di un'analisi attenta riguardo a ciò che può essere mantenuto e quello che deve essere cambiato o modificato necessariamente.

In seguito alle osservazioni riportate, pertanto, non risulta difficile definire il post-editing come una pratica che differisce dalla revisione in vari aspetti, e che può essere considerata, grazie alle sue caratteristiche e alle abilità richieste, come una particolare forma di traduzione.

### **2.1.2 ISO 18587 – Post-Editing Standard**

Lo standard ISO 18587 “Translation services - Post-editing of machine translation output - Requirements”, pubblicato nel 2017, è lo standard interamente dedicato al post-editing. Lo standard è da applicarsi solamente ai processi di post-editing, non si applica a processi di traduzione generali o di sistemi di traduzione automatica in generale.

Nella sua introduzione, lo standard ISO 18587 spiega le ragioni di adozione del PE:

The use of machine translation (MT) systems to meet the needs of an increasingly demanding translation and localization industry has been gaining ground. Many translation service providers (TSPs) and clients have come to realize that the use of such systems is a viable solution for translating projects that need to be completed within a very tight time frame and/or with a reduced budget<sup>5</sup>. (ISO 18587, 2017:V)

I vantaggi nell'uso della MT, dunque, sono principalmente volti all'incremento della produttività, grazie anche alla riduzione delle tempistiche di traduzione; il PE permette al traduttore, inoltre, di rimanere competitivo in un ambiente globalizzato in cui i clienti mostrano un incremento della domanda nell'uso dei sistemi di traduzione automatica all'interno delle traduzioni (ISO 18587, 2017:V). Il PE, come indicato sopra, è particolarmente indicato nei casi in cui vi sia un budget disponibile limitato: infatti

---

<sup>5</sup> L'uso di sistemi di traduzione automatica (MT o TA) ha preso piede per andare incontro alle necessità dell'industria di traduzione e localizzazione, le cui richieste sono in continuo aumento. Molti fornitori di servizi di traduzione (TSPs) e clienti hanno realizzato che l'uso di tali sistemi è una valida soluzione per i progetti di traduzione che devono essere completati in un lasso di tempo molto ristretto o con un budget ridotto (tradotto da me).

permetterebbe un abbassamento dei costi della traduzione. Nell'introduzione allo standard, viene specificato che non esiste un sistema di traduzione automatica che possa produrre un output, un risultato, pari a quello che potrebbe essere prodotto dall'uomo, e, pertanto, la responsabilità della qualità della traduzione finale cadrà sulla figura del traduttore, in base alla sua competenza in ambito di post-editing (ibidem). Quanto affermato ha una grande importanza, poiché aiuta il traduttore stesso a capire cosa ci si possa aspettare dai sistemi di traduzione automatica e che una buona qualità della traduzione non è sempre assicurata. Il traduttore che si accinge a effettuare un lavoro di post-editing, pertanto, deve tener conto delle varie problematiche con cui si potrebbe scontrare.

Lo standard si sofferma sull'obiettivo riproposto dal post-editing: "Post-editing is performed on MT output for the purpose of checking its accuracy and comprehensibility, improving the text, making the text more readable, and correcting errors"<sup>6</sup> (ivi, p. 5). Viene sottolineato che il post-editing differisce dalla traduzione manuale in quanto coinvolge e comprende tre diversi testi: il testo sorgente (*source text*), l'output del sistema di traduzione automatica e il testo finale di arrivo (*target text*) (ibidem).

All'interno dello standard ISO 18587 è evidenziata l'esistenza di due livelli principali di post-editing: *light PE* e *full PE* (che verranno approfonditi in seguito, ai punti 2.2.1 e 2.2.2), il cui uso dipende dalla scopo dell'output della traduzione, quindi ciò che viene richiesto dal testo finale, e dalle esigenze del cliente (ibidem).

Inoltre, vengono chiariti alcuni concetti di base fondamentali per comprendere appieno il processo di post-editing e i soggetti coinvolti. Questi ultimi sono, dunque: il cliente, il *translation service provider* (TSP), il quale si occupa di fornire il servizio di traduzione, il project manager, che si dedica agli aspetti specifici della traduzione o del post-editing ed è responsabile del processo, e ovviamente il post-editor (ivi, 3-4). Sotto la dicitura di *translation service provider* rientrano sia un singolo soggetto che un insieme di soggetti. Si può dunque trattare di una società di traduzione, un'agenzia di traduzione, un'organizzazione (profit, non-profit o governativa), un singolo traduttore freelance o post-editor, o ancora un ufficio interno addetto alla traduzione (ivi, 4). Il

---

<sup>6</sup> Il post-editing viene eseguito sull'output della traduzione automatica con lo scopo di controllare la sua accuratezza e comprensibilità, migliorando il testo, rendendo il testo più leggibile, e correggendo gli errori (tradotto da me).

TSP, inoltre, occupa un ruolo di fondamentale importanza in quanto, secondo lo standard, è colui che deve determinare se il contenuto del testo sorgente è adatto al sistema di traduzione automatica di cui si intende far uso e di conseguenza se risulta adatto al post-editing (ivi, 5). L'efficienza del post-editing dipende dal tipo di sistema di traduzione automatica, dalle combinazioni linguistiche, dal dominio e dal tipo di testo che si va a tradurre. Il TSP deve inoltre comunicare specifiche rilevanti al post-editor, come chi sia il destinatario del testo o quale livello di qualità sia richiesto per il testo d'arrivo, oltre a dover controllare la correttezza del formato del testo sorgente, affinché il post-editor non abbia problemi ad accedervi e consultarlo (ivi, 5-6).

Lo standard ISO 18587 fornisce una serie di scopi che devono essere perseguiti durante il processo di post-editing (ivi, 6):

- l'output del post-editing della traduzione automatica deve essere comprensibile;
- il contenuto del testo sorgente deve corrispondere al contenuto del testo d'arrivo;
- il post-editor deve attenersi alle richieste e alle specifiche precedentemente concordate.

Il testo finale deve risultare corretto sotto ogni punto di vista, ovvero dal punto di vista sintattico, della punteggiatura, dell'ortografia, della formattazione, tutti gli aspetti, dunque, sui quali, normalmente, il TSP e il cliente si accordano in precedenza. Il controllo del testo finale spetta al TSP, il quale si assicura del corretto svolgimento del lavoro richiesto, per poi inviarlo al cliente. Il post-editor deve riferire al TSP eventuali problemi riscontrati con il sistema di traduzione automatica, fornendo riscontri sulla qualità dell'output da esso prodotto, affinché possano essere migliorate le sue prestazioni e, di conseguenza, le traduzioni prodotte (ivi, 7).

Il post-editor, in quanto tale, deve possedere una serie di competenze fondamentali, oltre a delle qualificazioni, ottenute tramite previa formazione sul campo. All'interno dello standard ISO 18587 vengono riportate alcune competenze fondamentali (ivi, 7-9):

- il post-editor deve possedere sia delle competenze traduttive che linguistiche e testuali in entrambe le lingue (di partenza e di arrivo);
- il post-editor deve saper condurre efficienti ricerche per trovare ed esaminare le informazioni di cui necessita;

- è essenziale che il post-editor possieda delle competenze di tipo culturale affinché sia garantita la comprensione del testo da parte del pubblico a cui esso è rivolto;
- il post-editor deve possedere delle competenze tecniche: deve, infatti, essere in grado di utilizzare gli strumenti necessari allo svolgimento del post-editing;
- il post-editor deve, infine, essere competente e aver conoscenza dell'ambito a cui il testo si riferisce.

Nella successiva sezione vengono analizzate nel dettaglio le diverse tipologie di post-editing.

## 2.2 Tipi di post-editing

All'interno dello standard ISO 18587 è riportata la distinzione tra due diverse tipologie di post-editing, da poter applicare su richiesta del cliente: “*light* post-editing” e “*full* post-editing”. Inoltre, vi sono delle ulteriori differenze su come può essere effettuato il post-editing: esso può essere, infatti, bilingue o monolingue.

In modo particolare, con i termini *light* e *full* post-editing, ci si riferisce a una serie di linee guida da seguire durante il post-editing. All'interno di questo testo verrà fatto riferimento precisamente alle linee guide suggerite dalla società TAUS (Translation Automation User Society), una società nata nel 2005 con l'obiettivo di innovare e automatizzare la traduzione (TAUS - [www.taus.net/company/about-us](http://www.taus.net/company/about-us)).

### 2.2.1 Light PE

Secondo lo standard ISO 18587 con il termine *light* PE si intende un processo di post-editing volto a ottenere un testo meramente comprensibile, senza alcun tentativo di produrre un prodotto comparabile a uno ottenuto attraverso la traduzione umana<sup>7</sup> (ISO 18587, 2017:2).

Conformemente a quanto espresso dalla società TAUS, il *light* post-editing è quella tipologia di post-editing che permette di arrivare a un risultato di una qualità *good enough*, ossia una traduzione che non necessita di essere perfetta, ma solamente comprensibile. Con questo si intende che il testo deve essere chiaro nei suoi contenuti, accurato (ovvero il senso del testo sorgente deve essere mantenuto), mentre la qualità stilistica passa, tuttavia, in secondo piano (TAUS, 2016:17). Il *light* PE è da considerarsi

---

<sup>7</sup> “The process of post-editing to obtain a merely comprehensible text without any attempt to produce a product comparable to a product obtained by human translation” (ISO 18587, 2017:2)

come una qualità di minor livello, detta anche *fit for purpose* (ivi, 16), cioè che è volta e adatta a un determinato scopo. Quando lo standard di traduzione prefissato è basso, dunque, la scelta migliore ricade sul *light PE*. Ciononostante, se la qualità dell'output del sistema di traduzione automatica è molto bassa, un *light PE* non sarà sufficiente a ottenere una traduzione che rispetti gli standard di qualità idonei alla pubblicazione; se invece la qualità del *raw output* è adeguata, vi sono buone probabilità che sia sufficiente applicare le linee guida del *light PE* al fine di ottenere una traduzione di livello appropriato e, di conseguenza, raggiungere un livello di qualità adatto alla pubblicazione (ivi, 16).

Per l'ottenimento di questo tipo di qualità vi sono alcune linee guida da seguire (ivi, 17):

- è necessario puntare a una traduzione semanticamente corretta;
- il post-editor si deve assicurare che non sia stata aggiunta o rimossa qualche informazione in modo accidentale;
- è essenziale modificare qualsiasi tipo di contenuto che possa essere ritenuto offensivo, inappropriato o culturalmente inaccettabile;
- è fondamentale usare la maggior parte possibile di *raw output* fornito dal sistema di traduzione automatica;
- il post-editor deve applicare le regole basilari che riguardano l'ortografia;
- non vi è la necessità di applicare correzioni di sola natura stilistica;
- non risulta indispensabile la riformulazione delle frasi in modo solenne per migliorare il testo e la sua natura.

Una delle difficoltà maggiori che incontrano i traduttori nell'effettuare un *light post-editing* risiede proprio nel fatto di tralasciare i vari errori presenti nel testo, se non diversamente richiesto in modo esplicito dal cliente: bisogna attenersi alle linee guida, in quanto non si verrà pagati per le correzioni ulteriori, e nemmeno per il tempo usato per esse (Nitzke & Hansen-Schirra, 2021:31).

### **2.2.2 Full PE**

Secondo lo standard ISO 18587 con il termine "*full post-editing*" si intende un processo di post-editing volto a ottenere un prodotto comparabile a uno ottenuto da una

traduzione effettuata dall'uomo<sup>8</sup> (ISO 18587, 2017:2). L'obiettivo finale, dunque, risulta essere una traduzione che non può essere distinta da una effettuata dall'uomo, o il più possibile simile a essa. Se si vuole ottenere una traduzione di alto livello, che presenta, dunque, una qualità pari alla *human translation quality*, è raccomandato seguire le linee guida relative al *full post-editing* (TAUS, 2016:16).

Quando si parla di full PE si intende una qualità opposta a quella proposta dal *light PE*: infatti, nel caso del *full PE* il livello di accuratezza e comprensibilità dev'essere notevolmente alto, e la qualità stilistica molto buona. Ciononostante, potrebbe comunque risultare di qualità inferiore rispetto a una traduzione effettuata da zero da un traduttore professionista madrelingua (ivi, 18). Grande attenzione va dedicata alla grammatica, alla punteggiatura e alla sintassi.

Anche in questo caso la società TAUS fornisce una serie di indicazioni come linee guida per effettuare una traduzione con una qualità pari a quella umana (ibidem):

- lo scopo del *full PE* è quello di ottenere una traduzione corretta dal punto di vista grammaticale, sintattico e semantico;
- è necessario assicurarsi che la terminologia chiave sia correttamente tradotta. Può essere omessa dalla traduzione esclusivamente la terminologia appartenente a una lista fornita dal cliente, in cui sono inseriti i termini che non devono esplicitamente essere tradotti;
- è essenziale assicurarsi che nessuna informazione sia stata omessa o aggiunta accidentalmente;
- risulta fondamentale modificare qualsiasi tipo di contenuto che possa essere ritenuto offensivo, inappropriato o culturalmente inaccettabile;
- anche in questa casistica è indispensabile usare la maggior parte possibile di *raw output* fornito dal sistema di traduzione automatica;
- devono imprescindibilmente essere applicate le regole basilari riguardanti l'ortografia, la punteggiatura e l'*hyphenation*<sup>9</sup>;
- bisogna infine assicurarsi che la formattazione sia corretta.

---

<sup>8</sup> Full PE: “process of post-editing to obtain a product comparable to a product obtained by human translation” (testo originale ISO 18587)

<sup>9</sup> Con il termine *hyphenation* si intende l'unione di due parole attraverso l'uso di un trattino (ad esempio la stessa parola “post-editing”) (Cambridge Dictionary Online).

Il full PE permette di ottenere una traduzione molto più scorrevole e corretta sotto diversi punti; ciononostante, come già evidenziato per il caso del light PE, è necessario mantenere il più possibile di *raw output* per non spendere eccessivo tempo sul post-editing del testo (Nitzke et al., 2021: 32).

### **2.2.3 PE monolingue e bilingue**

Normalmente, il testo d'arrivo prodotto dal post-editor è frutto di una rielaborazione (se necessario) dell'output prodotto dal sistema di traduzione automatica, il quale viene confrontato con il testo sorgente. Questo processo viene anche chiamato “post-editing bilingue” e rappresenta la prassi tipica nella maggioranza delle situazioni: il post-editor confronta il testo di partenza e quello di arrivo, controllando la qualità della traduzione e assicurandosi che sia stata trasferita nel testo d'arrivo la giusta quantità di significato espresso nel testo sorgente (Nitzke & Hansen-Schirra, 2021:42).

Vi sono, tuttavia, dei casi in cui il post-editing può essere effettuato senza che venga preso in considerazione il testo sorgente, e, di conseguenza, risulta di fondamentale importanza che il traduttore possieda un'ottima padronanza linguistica solamente nella lingua target, di arrivo, mentre il ruolo della lingua di partenza passa in secondo piano, non avendo a disposizione il testo sorgente: si tratta del “*monolingual post-editing*” (MPE), ovvero il post-editing monolingue (ibidem). Come riporta Koponen (2016b:18), questo approccio venne suggerito già nel 1958 da Edmundson e Hayes, due dei primi autori ad aver scritto sul post-editing. Essi presumevano che il post-editor potesse servirsi, ai fini del suo lavoro, del testo tradotto automaticamente dalla MT e di un codice grammaticale, il quale indicava la parte del discorso, il caso, il numero e altri dettagli di ogni parola, senza la necessità di consultare il testo sorgente (ibidem).

Ma com'è possibile, per un traduttore, effettuare il post-editing di un testo, controllando dunque l'output prodotto da un sistema di traduzione automatica e apportando le modifiche e correzioni necessarie, senza disporre del testo di partenza, in lingua originale? Secondo Krings (2001:182-183, in Koponen, 2016b:18), vi sono alcune tipologie di errori, ad esempio, errori stilistici e/o grammaticali, che sono facilmente riconoscibili e di conseguenza correggibili senza servirsi del testo originale, mentre altri errori, quali, tra gli altri, quelli relativi all'incorrettezza di alcune parti del testo, risultano più difficilmente individuabili o correggibili. Gli errori tralasciati o impossibili da poter correggere senza il confronto con il testo originale, portano, in

svariati casi, a un cambiamento drastico del significato delle frasi componenti il testo (Kring, 2001:275, in Koponen, 2016b:18). Krings (2001:532) scoprì che il MPE portava con sé dei vantaggi: il post-editing, effettuato senza il confronto con il testo sorgente, risultava infatti un procedimento nettamente più rapido rispetto al PE bilingue, ma, considerate le problematiche che esso generava, non era da considerarsi come la migliore tra le opzioni (in Koponen, 2016b:18).

Koponen riporta due studi: il primo di Koehn (2010:541, in Koponen 2016b:18), e il secondo di Callison Burch *et al.* (2010, in Koponen, 2016b:138). Entrambe le ricerche sono volte a valutare l'efficacia (o meno) del post-editing monolingue e ambedue mostrano risultati che differiscono molto tra loro, a seconda della coppia linguistica presa in considerazione e in base al sistema di traduzione automatica utilizzato. Nello studio di Koehn (2010) vengono considerate come corrette con successo tra il 26% e il 35% delle frasi soggette a MPE; nell'analisi di Callison-Burch *et al.* (2010) i risultati mostrano delle grandi variazioni: nel peggiore dei casi, solo il 10% delle frasi sottoposte a MPE vengono considerate come corrette, mentre nel migliore dei casi, l'80% delle frasi risultano corrette con successo dopo il MPE.

In uno studio successivo, condotto da Mitchell, Roturier e O'Brien (2013) sui post degli utenti di alcuni forum online, precedentemente tradotti automaticamente da un sistema MT e in seguito sottoposti a MPE, i ricercatori hanno valutato la qualità, fluidità, comprensibilità e fedeltà del significato delle traduzioni, determinando se il post-editing monolingue avesse apportato un miglioramento o, al contrario, un abbassamento della qualità della traduzione prodotta dal sistema di traduzione automatica. I risultati hanno dimostrato che il post-editing monolingue è meno efficace del post-editing bilingue nel perfezionamento della fedeltà al significato originale nella coppia linguistica inglese-tedesco (il 43% delle frasi è considerata corretta dopo il MPE contro il 56% delle frasi considerate corrette, dopo essere state sottoposte a PE bilingue), ma leggermente più spesso (67% contro il 64% delle frasi) nella coppia linguistica inglese-francese. Per quanto riguarda il miglioramento della comprensione del testo, il MPE si è sempre rivelato meno efficace rispetto al PE bilingue: nella coppia linguistica inglese-tedesco, il 57% delle frasi è migliorato con il MPE rispetto al 64% delle frasi nella condizione bilingue. Per la coppia linguistica inglese-francese, il MPE rende migliore solo il 48% delle frasi, contro il 63% delle frasi migliorate nel PE

bilingue. In termini di fluidità, i post-editing monolingue e bilingue hanno avuto lo stesso successo nella coppia inglese-francese (63% delle frasi migliorate in entrambe le condizioni), e molto simile nella coppia inglese-tedesco, con una percentuale leggermente inferiore nelle frasi sottoposte a MPE (67% contro il 70% delle frasi).

In una ricerca più recente, compiuta da Macketanz *et al.* (2017), vengono confrontati i punti di forza e di debolezza di diversi sistemi di traduzione automatica (*rule-based*, *phrase-based*, *neural*). Nonostante le migliorie apportate nei nuovi sistemi di traduzione automatica, i risultati ottenuti dallo studio dimostrano che continuano a essere prodotte delle traduzioni ancora poco accurate, che non possono, dunque, essere considerate come prive di errori. Gli esiti della ricerca provano ulteriormente che, se non si ha a propria disposizione un output prodotto da un sistema di traduzione automatica che sia di buona qualità, la scelta di effettuare un MPE potrebbe rivelarsi azzardata e in alcuni casi totalmente errata.

### **2.3 Competenze**

Ogni traduttore, nello svolgimento del suo lavoro, mette in pratica una serie di competenze acquisite. Naturalmente, anche il post-editor deve padroneggiare delle abilità, in quanto, oltre a essere di base lui stesso un traduttore, normalmente possiede un'ulteriore formazione nell'ambito del post-editing, sviluppando determinate capacità e competenze aggiuntive.

Il tema delle competenze che i post-editor devono possedere è stato vastamente analizzato sia in ambito accademico che professionale (Rico & Torrejón, 2012:167). Lo stesso concetto di “competenza”, tuttavia, sembra sollevare delle controversie, in quanto gli autori utilizzano il termine in maniera differente, riferendosi sia all'acquisizione di capacità che al processo cognitivo della traduzione (*ivi*, 168).

Le competenze del post-editor si rivelano fondamentali nello svolgimento e nell'esecuzione di alcuni processi e compiti. All'interno di questi ultimi rientrano, ad esempio, i processi relativi alla traduzione automatica, che coinvolgono, ovvero, la lettura dell'output prodotto dai sistemi di traduzione automatica, e il confronto di questo con il testo sorgente (*ibidem*). Vi sono, poi, processi legati sia alla produzione che alla valutazione del testo target: i primi si riferiscono alla produzione di un nuovo testo, sia da elementi già presenti nel testo o tramite l'inserimento di nuovi elementi (Krings e Koby, 2001:314, in Rico & Torrejón, 2012:168); i secondi comprendono la valutazione,

positiva o negativa, dell'output della MT, giudicando, cioè, la qualità dell'output, comparata alle aspettative del cliente (Rico & Torrejón, 2012:168). I post-editor devono essere a conoscenza dei diversi mezzi di cui servirsi (come dizionari, testi paralleli, ecc.) per il miglioramento dell'accuratezza lessicale e fraseologica (ivi, 168-169), oltre a doversi occupare dei processi fisici di scrittura, se vi fosse la necessità di rimuovere/aggiungere elementi (ivi, 169).

Di seguito sono riportate le principali competenze che i post-editor dovrebbero possedere.

### **2.3.1 Competenze dei post-editor**

Il post-editing è una pratica complessa, che richiede particolare attenzione e responsabilità da parte del traduttore. Per questo motivo, nel corso degli anni sono stati messi a punto diversi modelli di riferimento sulle competenze richieste al post-editor, ad esempio i modelli PACTE (2003) e EMT (2009) (citati in Nitzke *et al.*, 2019:247-248). L'elenco di competenze proposto in questa sezione è tratto dal lavoro di Nitzke *et al.* (2019), basato sul "PACTE Translation Competence model" (2003), un modello che indica le varie competenze necessarie nell'ambito della traduzione, e sul modello di revisione delle competenze di Robert *et al.* (2017), che condividono alcune delle competenze richieste sia per i traduttori che per i post-editor (in Nitzke *et al.*, 2019:247-248).

Le competenze richieste nel post-editing possono essere suddivise principalmente in due tipologie: competenze chiave e sub-competenze sussidiarie (ivi, 248). Le competenze chiave sono le seguenti:

- competenza nella valutazione dei rischi: il post-editor deve essere in grado di stimare i rischi legati a un determinato testo e che possono derivare dalla sua traduzione (ibidem);
- competenza strategica: in seguito alla valutazione dei possibili rischi legati a un testo, il post-editor può decidere se procedere con un *full* o *light* post-editing per il lavoro di traduzione da compiere, o se usare solamente il sistema di traduzione automatica. Per le traduzioni ad alto rischio, potrebbe essere necessario applicare un ulteriore livello di controllo della qualità dell'output, oltre al *full* post-editing (ibidem);

- competenza di consulenza: dopo aver valutato i rischi e preso le decisioni strategiche adeguate, il post-editor contatta il cliente, informandolo sui rischi potenziali e sulle strategie di risoluzione dei problemi, e fornendo dei consigli a riguardo. Post-editor e cliente devono, sulla base delle informazioni ricevute, trovare un punto d'accordo (ibidem);
- competenza di servizio: il post-editor deve essere in grado di calcolare i prezzi in modo competente, consapevole e trasparente, considerando la qualità dell'output della MT e l'impegno richiesto dal post-editing (cfr. "Translation competence model by EMT Expert Group", 2009, in Nitzke *et al.*, 2019:248). Inoltre, questa specifica competenza include la conoscenza nell'uso dei sistemi CAT, di revisione e i sistemi di traduzione automatica da parte del post-editor: in sintesi, egli dovrebbe conoscere il mercato della traduzione, e avere capacità di negoziazione con il cliente, trovando un punto di incontro tra le proprie necessità e quelle del cliente, proponendo un'offerta appropriata e realistica (Nitzke *et al.*, 2019:248).

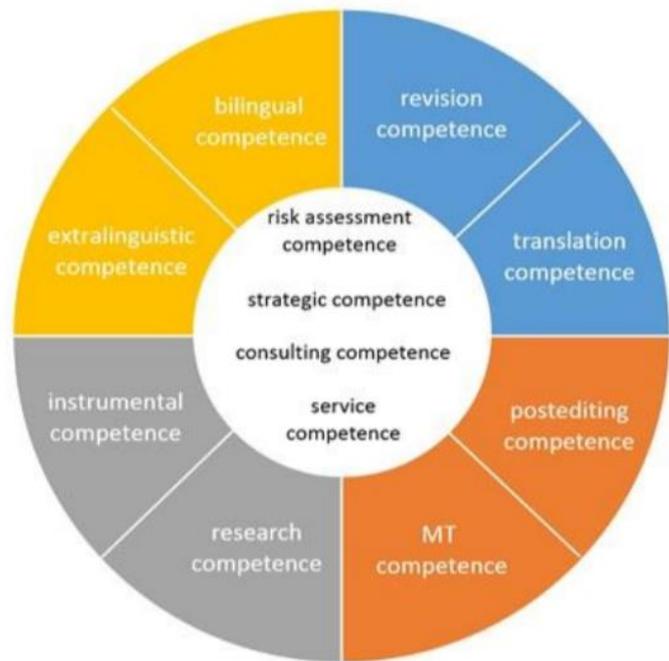
Vi sono, inoltre, una serie di sub-competenze, a supporto delle competenze chiave:

- competenza linguistica: comprende la conoscenza eccellente (livello madrelingua) sia della lingua di partenza che quella di arrivo (Rico & Torrejón, 2012:170; Nitzke *et al.*, 2019:249);
- competenze extralinguistiche: il post-editor deve conoscere la cultura di entrambe le lingue di lavoro, nonché una conoscenza generale del mondo e del dominio di appartenenza del testo sorgente, per poterne interpretare al meglio il suo significato (Nitzke *et al.*, 2019:249);
- competenza strumentale: il post-editor deve sapere come usare i CAT tool, al cui interno è integrato il sistema di traduzione automatica, e altre funzioni di cui il cliente potrebbe richiedere la conoscenza dell'uso (ibidem);
- competenza nella ricerca: il post-editor deve sapere come e dove trovare le informazioni che non conosce e/o di cui potrebbe avere bisogno (come ad esempio in dizionari specializzati, online, nei corpora, o ancora nei thesauri), in quanto un'efficace ed efficiente strategia di ricerca influenza la buona riuscita del lavoro di post-editing. Inoltre, il post-editor deve capire quanto

può fidarsi dell'output prodotto dal sistema di traduzione automatica (ibidem);

- competenza di revisione: il post-editor, analogamente alla revisione, deve sviluppare delle strategie affinché sia in grado di leggere attentamente un testo, scritto da un'altra persona (o prodotto da una MT). Egli deve, quindi, applicare delle strategie per la correzione di questi testi, individuando, oltretutto, la possibile presenza di errori significativi (ibidem);
- competenza traduttiva: il post-editing può essere definito come una particolare forma di traduzione. Dunque, il post-editor deve avere una buona conoscenza delle convenzioni delle diverse tipologie di testo, delle guide di stile, delle lingue controllate, delle specificità culturali, ecc. (ibidem);
- competenza nella traduzione automatica: il post-editor deve sapere come funziona un sistema di traduzione automatica e le insidie in esso nascoste, affinché diventi sempre più semplice l'individuazione delle problematiche. Inoltre, il post-editor deve essere in grado di valutare la qualità dei materiali di formazione e implementazione del sistema di traduzione automatica, o addirittura migliorare questo processo di formazione, se necessario (ivi, 250);

- competenza nel post-editing: i post-editor devono riuscire ad individuare gli errori presenti all'interno degli output prodotti dai sistemi di traduzione automatica neurale, che, essendo tecnologicamente più avanzati, producono un output scorrevole, ma che in realtà nasconde degli errori non ovvi e per questo più difficilmente



Figura

Nitzke *et al.*,

individuabili rispetto a quelli prodotti dai sistemi di traduzione assistita statistici (ibidem).

Il post-editing è influenzato anche da fattori circostanti, come riportato da PACTE (2003) e Robert *et al.* (2017) (in Nitzke *et al.*, 2019:251), quali le componenti psicofisiologiche, l'autopercezione del post-editor (alcune delle quali aiutano il post-editor nel miglioramento del livello di concentrazione e attenzione, nella resistenza allo stress, nel ragionamento logico, o ancora nel pensiero analitico), l'affinità nei confronti della tecnologia e del computer, il brief di PE che include le linee guida (il project manager dovrebbe racchiudere tutte le informazioni necessarie nel brief di post-editing, affinché la buona riuscita di quest'ultimo possa essere assicurata) (ibidem).

Nitzke *et al.*, (2019:251), parlando delle competenze necessarie per il PE, fanno un accenno al concetto di etica della traduzione (*translation ethics*), che dovrebbe essere presente di pari passo insieme alle competenze elencate. Il post-editor non dovrebbe percepire sé stesso come un semplice "correttore"/revisore dei testi prodotti da un sistema di traduzione automatica, ma come un competente esperto linguistico nella creazione dei processi di post-editing. Per questo, gli va riconosciuto il merito della

creazione del testo d'arrivo, sebbene quest'ultimo non risulti sempre perfetto: la capacità del post-editor risiede anche nell'accettare questa condizione di parziale imperfezione (Guerberof Arenas, 2013, in Nitzke *et al.*, 2019:251).

### **2.3.2 Formazione dei post-editor**

Nel corso degli ultimi anni, le richieste di post-editing della traduzione automatica (MTPE) hanno subito un grosso aumento, con la conseguente crescita di domanda, da parte dei servizi di localizzazione, di traduttori competenti in questo ambito e adeguatamente formati (Guerberof Arenas & Moorkens, 2019:217). In seguito alle nuove necessità del mercato della traduzione, diverse università hanno aggiunto dei programmi di formazione in campo del MTPE come parte dei loro curricula, e l'Universitat Autònoma de Barcelona (UAB) rappresenta il primo esempio di università ad avere introdotto l'insegnamento del post-editing e dei sistemi di traduzione automatica all'interno del programma del corso di laurea magistrale in Localizzazione (*ibidem*).

Come ricordato da Guerberof Arenas & Moorkens (2019:218), più di trent'anni fa lo scenario della formazione nell'ambito della traduzione era nettamente diverso, e gli studenti che sceglievano un percorso universitario orientato alla traduzione venivano preparati a tradurre libri, manuali, brochures, usando una macchina da scrivere elettronica, la cui invenzione, durante gli ultimi anni del XX secolo, fu una delle più grandi innovazioni introdotte nel mondo della traduzione, rivoluzionando il modo stesso di fare traduzione (*ibidem*). Rappresentò, infatti, un grande avanzamento tecnologico, e fu solamente il primo di una lunga serie. Con l'avvento del computer e la sua adozione in ogni agenzia di traduzione a partire dagli anni '90, e successivamente con l'arrivo di Trados Translator's Workbench (lanciato nel 1992, e già popolare tra moltissimi traduttori fin dal 1995), il lavoro del traduttore assunse un nuovo significato (*ibidem*). Grazie alla popolarità dei nuovi mezzi a disposizione, molte università introdussero, all'interno dei corsi di laurea triennale e magistrale, dei moduli in cui era previsto l'insegnamento dell'uso degli strumenti CAT (*ibidem*). Solamente alla fine degli anni '90 si cominciò a parlare di post-editing e la sua introduzione in grandi aziende (anche di sviluppo di software come Microsoft, IBM, Autodesk e SAP) avvenne a partire dai primi anni 2000: da questo momento iniziò la richiesta di figure adeguatamente

preparate a questa tipologia di lavoro, con la conseguente introduzione di programmi appositi a livello universitario (ibidem).

Diversi autori e ricercatori nel corso degli ultimi anni hanno proposto (e continuano a proporre) prototipi di corsi basati sull'insegnamento e la formazione nel campo del post-editing. Una prima proposta è quella di O'Brien (2002), che dichiara la necessità di una formazione separata dalla traduzione manuale e dalla revisione, richiedendo il post-editing abilità diverse da quelle richieste dalla traduzione. O'Brien fornisce, infatti, una serie di motivazioni a favore di una formazione nel campo del MTPE: il PE va incontro all'aumento di domanda di traduzione e velocizza i tempi di produzione; le capacità di un traduttore non sono le stesse di un post-editor e da ciò si può evincere che un eccellente traduttore, senza una formazione adeguata, non necessariamente sarà un buon post-editor; la formazione universitaria in ambito del post-editing permette di inserire nel mondo del lavoro delle persone già formate; inoltre, la formazione contribuisce a migliorare le percezioni del traduttore riguardo ai sistemi di traduzione automatica (O'Brien, 2002:100). O'Brien (2002) offre la sua proposta di un corso basato sul post-editing e sulla MT, riassumendo le competenze da acquisire necessariamente e le applicazioni pratiche del post-editing. Fondamentale, secondo la sua visione, è anche l'insegnamento delle basi della programmazione e della linguistica computazionale, oltre all'affinamento della padronanza degli strumenti CAT, con la creazione di memorie di traduzione, che possono fungere già da base personale per lo studente che, dopo aver terminato gli studi, diverrà traduttore qualificato (ibidem).

Kontinen *et al.* (2021:189), citano diversi autori che si sono interessati alla formazione dei post-editor, come Doherty e Kenny (2014), Flanagan e Christensen (2014), o ancora Koponen (2015), i quali propongono idee di corsi universitari in cui gli studenti vengono formati sulle basi della MT e del PE, attraverso l'applicazione pratica degli insegnamenti teorici (ad esempio, seguendo le linee guida della società TAUS per *light* e *full* post-editing). Secondo Kontinen *et al.* (2021), risulta controproducente separare il post-editing e la revisione durante la formazione universitaria: infatti, distaccando revisione e post-editing, gli studenti potranno sviluppare migliori capacità in una delle due attività, ma le due pratiche verranno considerate come totalmente diverse tra di loro, escludendo una o l'altra (Kontinen *et al.*, 2021:195-196). La loro proposta, dunque, consiste nel far coesistere i due insegnamenti e affinare entrambe le

pratiche, in quanto esse, se da una parte presentano grosse differenze, dall'altra hanno diversi punti in comune, e, attraverso un'analisi parallela, è possibile evidenziarne le somiglianze e le differenze.

Ginovart Cid e Colominas Ventura (2021) hanno compiuto uno studio sui programmi di formazione dei post-editor, sottoponendo alcune questioni relative al PE a 53 educatori nel campo, provenienti da 17 paesi diversi. Il focus è stato posto su tre principali aspetti, ovvero gli elementi della formazione sul post-editing, le abilità nel post-editing e i compiti del post-editing (ivi, 228-229). Dallo studio si evince che molti educatori nel campo non includono l'addestramento del sistema di traduzione automatica o la sua personalizzazione all'interno dei programmi post-laurea europei (ivi, 244). Oltretutto, sono pochi i docenti che inseriscono la pratica del MTPE all'interno dei corsi di traduzione (ibidem). Il problema evidenziato da un paio di intervistati sta proprio nel giudicare il post-editing come una pratica a sé stante, e non come un'ulteriore risorsa e sviluppo della traduzione (ivi, 245).

La ricerca di Stasimioti e Sosoni (2019:134), volta a identificare se effettivamente la formazione possa giocare un ruolo cruciale nel lavoro di post-editing, dimostra che, con un'adeguata preparazione, la percezione degli stessi studenti nei confronti del post-editing assume un ruolo diverso: gli studenti hanno un atteggiamento più ottimista e, inoltre, sono in grado di effettuare il post-editing di un testo più velocemente, evitando casi di *over-editing* (ibidem).

Risulta, quindi, indispensabile, in un panorama sempre più altamente specializzato, e in cui la concorrenza pervade il mondo del lavoro, possedere delle certificazioni ulteriori che dimostrino la propria abilità nell'uso di programmi di traduzione automatica e post-editing, in quanto molto spesso viene richiesto. Per questo motivo, è anche essenziale che le università forniscano una preparazione adeguata a coloro che vogliono proseguire il loro percorso nell'ambito del post-editing.

#### **2.4 Valutazione del post-editing**

L'output prodotto da un sistema MT, nonostante i continui progressi e le numerose modifiche che vengono costantemente apportate per migliorarne la qualità, non può essere considerato perfetto, e continua a presentare degli errori, dovuti ai sistemi di traduzione automatica stessi. Vi sono una serie di errori che si ripetono sistematicamente, o quasi, in tutti i sistemi di traduzione: ad esempio, nei sistemi di

traduzione statistici SMT non è raro incorrere in errori ortografici, grammaticali, di sintassi, ecc. In tale casistica, queste inesattezze risultano facilmente individuabili. Il vero problema si pone quando si tratta della traduzione automatica neurale, il cui output è apparentemente perfetto, in quanto lineare e leggibile in modo scorrevole: qui si trovano gli errori che più facilmente possono sfuggire all'occhio umano.

Legati agli errori prodotti dai sistemi di traduzione automatica vi sono due concetti chiave: qualità e produttività. Si tratta di due nozioni tra di loro strettamente collegate, poiché più la qualità dell'output prodotto dalla MT è buona, minore sarà l'intervento del post-editor e maggiore sarà la produttività del post-editor, portando a un abbattimento di tempo e costi.

Vengono di seguito analizzati nel dettaglio gli errori tipici dei sistemi di traduzione automatica, confrontando gli errori specifici dei sistemi di traduzione statistici (SMT) e neurali (NMT). Inoltre, viene esaminato il significato dei termini “qualità” e “produttività”, osservando la relazione esistente tra di loro.

#### **2.4.1 Errori tipici dei sistemi di traduzione automatica neurale**

I sistemi di traduzione automatica, anche i più evoluti e precisi, commettono degli errori che si ripetono in modo quasi sistematico, e che normalmente non verrebbero compiuti dai traduttori umani. Gli errori possono essere di diverso tipo, e secondo lo standard ISO 18587 sono suddivisi in errori stilistici, grammaticali, di traduzione letterale, di traduzione di nomi/denominazioni che dovrebbero rimanere invariati (ISO 18587, 2017:Annex A).

I sistemi di traduzione neurale commettono meno errori rispetto ai sistemi di traduzione statistici (SMT), le quali inesattezza consistono soprattutto in errori di ortografia, in uno scorretto uso delle maiuscole, nella mancanza di coerenza, nell'errata traduzione del contenuto, errori di sintassi, omissioni o aggiunte non richieste, ecc. (Arnhold *et al*, 2017:236-238). Tuttavia, l'output ottenuto dalla traduzione neurale è ancora lontano dall'essere considerato perfetto.

I sistemi di traduzione automatica neurale (NMT) presentano una qualità inferiore al di fuori del dominio<sup>10</sup>, al punto da sacrificare completamente l'adeguatezza a favore della scorrevolezza del testo: dunque, le parole all'interno di domini differenti, vengono

---

<sup>10</sup> Un dominio è definito da un corpus proveniente da una fonte specifica e può differire da altri domini per argomento, genere, stile, livello di formalità, ecc. (tradotto da Koehn e Knowles, 2017:29).

tradotte in modi diversi, e conseguentemente anche il significato deve essere espresso in uno stile diverso (Koehn e Knowles, 2017:29). Per questo motivo devono essere sviluppati dei metodi per l'adattamento del dominio, allenando il sistema di traduzione automatica attraverso l'inserimento di dati (ibidem). Lo studio compiuto da Koehn e Knowles (2017:29-30), mostra che, se fuori dal proprio dominio, la traduzione con sistemi NMT risulta addirittura peggiore di quella effettuata con sistemi SMT: l'output del sistema NMT appare fluido, ma sconnesso dall'input iniziale (ivi, 30).

I sistemi NMT hanno una curva di apprendimento più ripida rispetto alla quantità di dati di addestramento, mostrando una qualità peggiore in contesti con basse risorse, ma prestazioni migliori in contesti ad alte risorse: la performance è migliore quando si usano grandi quantità di dati rispetto a quando ve ne sono pochi (ivi, 31).

Normalmente si pensa che i modelli di NMT non siano particolarmente performanti per quanto riguarda la traduzione di parole a “bassa frequenza”, parole rare, in particolare quelle appartenenti a categorie altamente flesse (come, ad esempio, i verbi o gli aggettivi): questo è uno degli errori che ancora devono essere risolti nei sistemi NMT, in quanto queste sono parole poco allenate o addirittura nuove per il sistema di traduzione automatica (ibidem). Si tratta in modo particolare di nomi propri di persona (o cognomi), i quali potrebbero non essere mai stati esaminati dal sistema fino a quel momento, oppure ancora nomi di luoghi, località meno conosciuti (ivi, 33).

Un'altra problematica risiede nelle frasi particolarmente lunghe: i sistemi NMT hanno una qualità di traduzione inferiore su frasi molto lunghe e funzionano meglio su frasi di circa 60 parole (ibidem). Inoltre, non sempre le parole, una volta tradotte, vengono allineate nel modo corretto: è stato infatti imposto un modello (*attention model*) di allineamento dell'output rispetto all'input (ivi, 34). Vi è la necessità, infatti, di un meccanismo di allineamento tra le parole del testo sorgente e di quello di arrivo (ibidem). Infine, Koehn e Knowles citano il *beam search*, un algoritmo che utilizza una ricerca in ampiezza: il compito della decodifica è quello di trovare la traduzione completa della frase con la massima probabilità (ivi, 35).

Se si sottopone l'output prodotto dai sistemi NMT alla valutazione umana, si nota, come riportato nello studio di Castilho *et al.* (2017:114), che questi sistemi continuano ancora a presentare molti errori di omissione di parole, risultando meno performanti dei sistemi SMT in alcuni casi (nello studio in questione solo il 2% dei segmenti tradotti

con sistemi NMT era privo di errori contro il 25% dei segmenti tradotti con sistemi SMT).

Risulta necessario sottolineare che i sistemi di traduzione automatica sono in continuo sviluppo e addestramento e l'andamento va verso sempre una maggiore perfezione di questi sistemi: nel periodo storico in cui ci troviamo giornalmente vengono apportate migliorie ai sistemi di traduzione neurale. Studi attuali rivelano come ci si stia avvicinando sempre più alla *singularity*, ovvero, come spiega il CEO di Translated Marco Trombetti, in un'intervista per Slator, quando un traduttore professionista umano impiega meno tempo a editare una traduzione effettuata da un sistema MT che non revisionare una traduzione effettuata da un altro traduttore professionista umano (Slator.com<sup>11</sup>). Questo è possibile solamente se l'output prodotto dal sistema MT è talmente buono da non dover quasi nemmeno essere rivisto (vedi § 2.7 **Prospettive future** per un maggiore approfondimento sul tema della *singularity*).

#### 2.4.2 Produttività e qualità

Il termine “produttività” si riferisce a quanto velocemente una traduzione può essere effettuata (O'Brien, 2021:178). La produttività nel MTPE riguarda il numero di parole di cui il traduttore è in grado di svolgere il post-editing nel corso di un giorno, avendo precedentemente fissato il livello di qualità della traduzione finale con il committente (ibidem). Attorno alla produttività della traduzione, intesa nel suo senso più ampio, vi sono delle aspettative, che sono determinate da una serie di fattori, quali la combinazione linguistica, il tipo di contenuto e quindi la tipologia di testo presa in esame, il supporto tecnologico di cui si serve il traduttore (ibidem). Normalmente, un traduttore professionista produce tra le 2000 e le 3000 parole in un giorno, senza servirsi di *CAT tool* (ibidem); se vi è l'utilizzo di sistemi di traduzione automatica, quest'aspettativa si alza, in quanto si suppone che il post-editor sia molto più produttivo di un normale traduttore (ibidem). Tuttavia, non di rado le aspettative si spingono troppo oltre quella che è la realtà, dal momento che la produttività è strettamente legata alla qualità dell'output prodotto dal sistema di traduzione automatica. Sono diversi gli studi presenti sulla tematica, tra cui quello di Autodesk (2011), riportato da Guerberof Arenas (2014:168), in cui viene affermato che in generale la produttività data dal post-editing è maggiore rispetto a quella che si può ottenere attraverso la traduzione umana. La

---

<sup>11</sup> <https://slator.com/translated-marco-trombetti-time-to-edit-as-proxy-for-ai-singularity/>

produttività varia a seconda della combinazione linguistica (nello studio si sottolinea che la crescita della produttività è pari al 131% per la lingua francese, mentre solamente del 42% per la lingua cinese). L'esperienza nel post-editing, inoltre, sembra essere il fattore cruciale che denota la differenza nella produttività (ibidem). Altri studi, come quelli di Carl *et al.* (2011) e García (2010) (citati in Koponen, 2016a:136), non hanno dimostrato grandi variazioni tra la produttività ottenuta tramite il post-editing della traduzione automatica e la traduzione manuale. Lo studio di García (2010), tuttavia, coinvolgeva dei traduttori apprendisti, e non professionisti, e l'esperienza, come precedentemente sottolineato, si rivela un fattore cruciale nella pratica di post-editing (ibidem). Le ricerche sulla produttività, dunque, forniscono dei pareri tra loro discordanti. Tuttavia, O'Brien (2021:178) suggerisce che non è da considerarsi irrealistica la possibilità di riuscire ad aumentare almeno del 30% la propria produttività quando ci si occupa di post-editing: se si usa un buon sistema di traduzione automatica, allenato a lungo con dati di alta qualità, e se si ha notevole grado di preparazione e molta esperienza all'interno del campo, il post-editing può rivelarsi più veloce della traduzione manuale o effettuata tramite CAT *tool* (ivi, 178-179).

La produttività è strettamente legata alla qualità. La qualità prodotta dai sistemi di traduzione automatica negli ultimi due decenni è migliorata notevolmente e continua a progredire, grazie all'adozione dell'approccio statistico e neurale alla traduzione automatica (NMT e SMT) (ivi, 179). Tuttavia, la qualità dell'output della MT rappresenta uno degli aspetti sui quali i traduttori mostrano più diffidenza quando si approcciano al post-editing: essi temono che il prodotto finale si riveli di una qualità molto inferiore rispetto a quello che si potrebbe ottenere con una traduzione effettuata senza l'aiuto di una MT. Il timore risiede, infatti, nel danno che potrebbe essere inflitto alla reputazione del traduttore stesso (Arnhold *et al.*, 2017:233). Spesso i traduttori si trovano di fronte alla necessità di cancellare la traduzione suggerita e riscriverla da capo, in quanto la qualità dell'output della MT è troppo scarsa. Come già segnalato, sebbene i sistemi di traduzione automatica neurale abbiano apportato un netto miglioramento rispetto ai sistemi di traduzione automatica statistici (SMT), sono presenti ancora errori di diverso tipo, che, se richiesto, devono essere controllati dal post-editor (O'Brien, 2021:179). Uno dei problemi principali, risiede nel fatto che l'output prodotto dai sistemi di traduzione neurale, sebbene non sia sempre totalmente

corretto e, al contrario, vi siano casi di omissioni di intere sezioni di frasi, risulta scorrevole: gli errori sono, pertanto, meno visibili anche all'occhio di un traduttore esperto (ibidem). I sistemi SMT, invece, commettono degli errori che normalmente non verrebbero compiuti da un traduttore umano esperto e sono facilmente notabili durante il post-editing (ibidem).

La qualità del post-editing, però, può essere di diverso tipo, come visto in precedenza, a seconda del livello di post-editing che viene richiesto (*light* o *full* PE). Il livello della qualità deve essere rispettato, e, ovviamente, in base a esso si otterrà una traduzione più o meno accurata. Bisogna, quindi, fare molta attenzione all'incarico ricevuto: i testi post-editati non dovrebbero mirare alla perfezione linguistica bensì alla "frugalità" linguistica (ovvero il grado di soddisfazione che si ha con quella qualità) (Arnhold *et al.*, 2017:233).

Come segnalato in uno studio di Cadwell *et al.* (2018), diversi fattori influenzano la qualità dell'output del sistema di traduzione automatica. Essi sono rappresentati, ad esempio, dal tipo di sistema MT usato, dalla quantità di dati che sono stati allenati, dalla combinazione linguistica presa in considerazione, dalla qualità del testo sorgente (Cadwell *et al.*, 2018:313). In tale studio viene proposto anche un paragone tra l'output prodotto da un sistema di traduzione automatica e quello ottenuto attraverso le memorie di traduzione: viene evidenziato che sia i segmenti TM sia l'output MT possono essere di bassa qualità. Tuttavia, mentre i segmenti nei *translation memory systems* sono meno consistenti nel tipo di errori che possono contenere e le differenze tra il testo sorgente e quello di arrivo sono evidenziate, nell'output del sistema MT gli errori sono imprevedibili, inconsistenti e, conseguentemente, abbassano il livello di fiducia (ivi, 314). La fiducia gioca un ruolo fondamentale: è un elemento chiave che spinge o, al contrario, respinge i traduttori nell'approcciarsi al post-editing (ibidem). I traduttori, inoltre, ripongono maggiore fiducia dei suggerimenti proposti dalle TM, poiché è il traduttore stesso che crea le TM, e non una macchina esterna (ivi, 315).

Koponen (2016a) riporta uno studio di Fieder e O'Brien (2009:62-63), in cui viene qualificata la qualità dell'output in base a tre variabili: chiarezza (comprensibilità del testo), accuratezza (stesso significato della frase nel testo sorgente) e stile (appropriatezza allo scopo, naturalezza e attinenza idiomatica) (in Koponen, 2016a:137). Le traduzioni considerate all'interno della ricerca in questione, e sottoposte

a post-editing, vengono classificate come migliori per quanto riguarda la chiarezza e l'accuratezza, ma peccano di stile: gli intervistati dichiarano di preferire le versioni tradotte manualmente a quelle sottoposte a post-editing, dando più importanza allo stile (ibidem).

I concetti di qualità e di produttività sono a loro volta direttamente legati al tempo, e di conseguenza al costo del post-editing: più la qualità dell'output del sistema di traduzione automatica è alta, più si alzerà la produttività del traduttore, che riuscirà, dunque, a effettuare il post-editing di un numero di parole maggiore.

## **2.5 Post-editing nei translation memory systems**

La maggior parte dei traduttori professionisti, al giorno d'oggi, si serve dei sistemi CAT nel proprio lavoro di traduzione. Sebbene i sistemi CAT siano i privilegiati, rispetto ai sistemi di traduzione automatica, la richiesta di post-editing della traduzione automatica è in continuo aumento. L'uso di uno dei due sistemi non esclude, tuttavia, l'altro: chi si occupa di post-editing continua, in molti casi, a servirsi anche dei *CAT tool* nella propria routine lavorativa, in base alla tipologia del lavoro richiesto. Due delle tecnologie predominanti di cui si servono i traduttori professionisti per aumentare la loro produttività sono, per l'appunto, i sistemi di traduzione automatica e i *CAT tool* (Ortega *et al.*, 2019:256). Spesso ai traduttori, di fronte a queste due possibilità offerte dal mercato della traduzione, sorge spontanea una domanda: come è possibile inserire, nel proprio lavoro di tutti i giorni, i sistemi di traduzione automatica? La soluzione consiste nell'integrazione dell'output prodotto dal sistema di traduzione automatica all'interno dei *translation memory systems*. I sistemi differiscono tra di loro, ma i processi tramite i quali avviene questa integrazione sono molto simili, o sono addirittura uguali.

TM e MT sono spesso stati considerati come tecnologie completamente separate, ciò dovuto anche al loro emergere con tempistiche diverse nel mercato della traduzione (O'Brien, 2021:181). La tecnologia è in continua evoluzione, e i sistemi di traduzione automatica vengono integrati nei *translation memory systems*. Se abilitate, vi sono delle funzionalità particolari che permettono di sostituire, all'interno di uno strumento CAT, un match inferiore al 70/75% con la versione della frase prodotta da un sistema di traduzione automatica (ibidem). Può anche essere decisa la soglia sotto la quale i suggerimenti della traduzione automatica possono sostituirsi ai *fuzzy match*, oppure

ancora usare l'output MT come un'opzione aggiuntiva, inserendola nel segmento ed effettuando il post-editing di esso: una volta confermato il segmento verrà aggiunto alla TM (Nitzke & Hansen-Schirra, 2021:45). Il prodotto finale di un segmento, dunque, può essere un *fuzzy match* editato dal traduttore, oppure può essere il risultato del post-editing dell'output di un sistema di traduzione automatica: si crea così un unico ambiente di lavoro in cui entrano in simbiosi tra loro funzioni diverse (O'Brien, 2021:181). Anche un *fuzzy match* può essere "sistemato" tramite un sistema di traduzione automatica: vi è una funzione avanzata chiamata *fuzzy match repair* che perfeziona automaticamente o revisiona il *fuzzy match*, usando le informazioni dalla TM, da un glossario o da un motore di un sistema di traduzione automatica (ibidem).

Già nel 2010 due ricercatori, Kanavos e Kartsaklis (2010) condussero uno studio a dimostrazione di come l'uso, da una parte di testi già tradotti attraverso le TM, e dall'altra tradurre nuove porzioni di testo con l'aiuto di un sistema di traduzione automatica, potesse rappresentare un grande vantaggio. Provarono che l'utilizzo in simbiosi dei due sistemi portò a un aumento della produttività alquanto significativo. In particolare, evidenziarono un aumento di produttività nei segmenti con una corrispondenza inferiore al 75-80% con le frasi estratte dalla TM. La loro ricerca è da considerarsi lungimirante: oggi molti *CAT tool* consentono l'integrazione dell'output prodotto da un sistema di traduzione automatica.

Normalmente un lavoro di post-editing può essere elaborato utilizzando da una parte i classici strumenti CAT (come, ad esempio, SDL Trados Studio, Across o Transit NXT) oppure dei *CAT tool* online di aziende o istituti di ricerca/università appositamente sviluppati per il post-editing (quali MateCat, CATaLog e altri) (Arnhold *et al.*, 2017:235). Gli strumenti CAT classici vengono acquistati dal provider e installati nel computer, permettendo al traduttore di lavorare anche offline: questa metodologia di lavoro è più sicura, poiché permette il salvataggio del lavoro anche nel caso di problemi con la connessione internet (ibidem). I *CAT tool* online sono, di norma, gratuiti, e portano con sé dei vantaggi, come il loro utilizzo ovunque e in qualunque momento. Il loro funzionamento esclusivamente online, tuttavia può comportare la perdita dei dati e del lavoro svolto, in caso di problemi di connessione (ibidem). Inoltre, questa tipologia di sistemi possiede solo delle funzioni basilari, come il controllo ortografico, i tags e la punteggiatura (ivi, 236). Trados Studio 2021/2022 è un esempio di *CAT tool* ibrido,

ovvero che può essere utilizzato sia in cloud che in locale, permettendo, dunque, di lavorare sia online che no, fornendo una sicurezza aggiuntiva ai suoi utilizzatori.

Il sistema di traduzione automatico dev'essere integrato o attivato/disattivato nel rispettivo sistema TM: può essere attivato immediatamente (e in questo scenario è necessario assicurarsi che il sistema di traduzione in questione sia adatto al progetto a cui si deve lavorare), o se è necessario attivare la componente MT manualmente (Nitzke & Hansen-Schirra, 2021:45). Una volta scelto il sistema di traduzione automatico per il proprio progetto, l'output della MT spesso viene inserito automaticamente nei segmenti vuoti (ovvero che non presentano *full* o *fuzzy match* presi dalla TM) (ibidem).

All'interno di alcuni TM *system* vi è la possibilità di utilizzare un sistema di traduzione automatica adattivo (“adaptive machine translation”); inoltre, la “*interactive* MT” è una funzione inclusa in alcuni strumenti, e comprende la modifica del suggerimento da parte del sistema di traduzione automatica in tempo reale (O'Brien, 2021:181). Per questa funzione viene messa in discussione la validità della dicitura “post-editing”, in quanto il processo avviene nell'immediato (ibidem).

Una critica che alcuni traduttori fanno nei confronti dei sistemi di traduzione automatica, confrontandoli con i TM *system*, consiste nel fatto che tanto i segmenti TM quanto l'output della MT possono peccare di qualità (Cadwell *et al*, 2018:314. Tuttavia, mentre i sistemi di traduzione automatica producono degli errori che sono imprevedibili, discordanti, scoraggiando la fiducia del post-editor, nei TM *system* gli errori sono meno consistenti nel tipo di errore che i segmenti possono contenere e nei TM *match* le differenze tra il testo source e quello target vengono sottolineate (ibidem). È fondamentale, pertanto, essere consapevoli dei rischi che si incorrono, alcune volte legati anche alla mancanza di consapevolezza riguardo all'origine dei suggerimenti della MT, che non di rado contengono inesattezze.

L'idea, al giorno d'oggi, è quella di creare sempre più una sinergia tra le due tecnologie in questione (Moorkens & O'Brien, 2017:109). Viviamo in un momento storico in cui le tecnologie di cui il traduttore si serve continuano a migliorare e a evolversi continuamente, e vi saranno nuove funzioni per far interagire in modo ancora più efficiente le TM e i sistemi di traduzione automatica.

## **2.6 Rischi e riservatezza dei dati nel post-editing**

Il traduttore ricopre un ruolo che racchiude in sé alcune insidie: sulla sua figura gravano, infatti, una serie di responsabilità. La traduzione, in generale, è un lavoro che deve essere svolto in maniera precisa, sotto tutti i suoi aspetti, perché può essa stessa essere considerata come un processo “rischioso”, ovvero che nel quale è facile incorrere in conseguenze importanti, anche a livello legale. È fondamentale, dunque, essere a conoscenza dei possibili rischi in cui è possibile imbattersi, sia durante il processo di traduzione che di post-editing: ad esempio, quando ci si occupa del post-editing dell’output prodotto da un sistema di traduzione automatica, è necessario ricordare che questo raramente è completamente perfetto, sebbene lo possa apparentemente sembrare. Effettuare il post-editing in maniera poco accurata potrebbe comportare dei rischi alla qualità del prodotto finale, compromettendo la reputazione stessa del post-editor.

Un ulteriore aspetto su cui risulta cruciale porre la propria attenzione è la protezione dei dati degli utenti e clienti, che vengono inseriti all’interno di sistemi di traduzione automatica. Non tutti i sistemi di traduzione consentono una totale protezione dei dati personali: per questa ragione sarebbe doveroso controllare la politica sul trattamento dei propri dati per ogni sistema che si va a utilizzare.

Nei seguenti punti viene proposta un’analisi dei principali rischi legati al post-editing e alcune fondamentali informazioni riguardanti la regolamentazione sulla protezione dei dati personali, inseriti all’interno di sistemi di traduzione automatica.

### **2.6.1 Rischi del post-editing**

La traduzione comporta dei rischi per tutti gli attori coinvolti nel suo processo (Canfora & Ottmann, 2020b:136). Le decisioni che vengono prese durante l'intero processo di traduzione sono sostenute dagli stessi principi che regolano le decisioni prese a qualsiasi altro livello aziendale (Nitzke *et al.*, 2019:240). Lo strumento per lo sviluppo di criteri decisionali in economia è rappresentato dalla gestione del rischio (*risk management*) (ibidem). In ambito economico e aziendale, i criteri decisionali si dividono in decisioni strategiche (a lungo termine), tattiche (a medio termine) e operative (a breve termine) (Hofmann, 2012:173, in Nitzke *et al.*, 2019:240). Quando si considera la gestione del rischio per le situazioni in cui il post-editing è coinvolto, le decisioni aziendali sono di due tipologie: strategiche e operative (Nitzke *et al.*, 2019:240). Le decisioni strategiche riguardano, ad esempio, la volontà o meno di utilizzare i sistemi di traduzione automatica da parte dell’organizzazione, mentre le decisioni operative si riferiscono, ad

esempio, alla scelta delle linee guida considerate necessarie per il testo preso in considerazione, tenendo conto delle decisioni strategiche dell'organizzazione (ibidem).

La gestione del rischio è regolata dallo standard internazionale ISO 31000 “Risk management – Principles and guidelines” (2018): si tratta di uno standard orizzontale, e per questo motivo, sebbene non sia esplicitamente e solamente riferito alla traduzione, può essere applicato al processo di traduzione e al post-editing (ibidem). Vi sono una serie di rischi che emergono dai processi di post-editing (Nitzke *et al.*, 2019:241):

- violazione dei dati: si tratta della diffusione via web di informazioni confidenziali che si trovano in un sistema di traduzione automatica *web-based* (ibidem);
- perdita di controllo dei processi: il cliente non può controllare se il traduttore usa in modo trasparente il sistema di traduzione automatica o se le sue funzionalità sono trasparenti per l'utente del sistema di traduzione automatica, specialmente nel caso della traduzione neurale (ibidem);
- aumento della responsabilità per il cliente: questo rischio si presenta quando le istruzioni per l'uso sono tradotte tramite il post-editing della traduzione automatica, in particolare nel caso di testi considerati come ad alto rischio (ibidem);
- atteggiamento nei confronti della MT: i clienti potrebbero incontrare difficoltà nel trovare dei traduttori e post-editor qualificati, poiché molti traduttori professionisti hanno dei pregiudizi nei confronti del post-editing (ibidem);
- rischi nella qualità: la qualità del testo dopo il post-editing potrebbe non essere sufficientemente adeguata ai bisogni del cliente o del gruppo target (ibidem).

Canfora e Ottmann (2020a), forniscono delle ulteriori informazioni riguardo i rischi possibili che si incontrano in particolare nella traduzione automatica neurale, dividendoli in tre aree di rischio: possibili danni verso i clienti, problemi di responsabilità e *cyber-risks*, ovvero i rischi informatici, derivanti dalle attività compiute in rete (tra cui rientra il problema della protezione dei dati, che diventano liberamente accessibili una volta inseriti online) (Canfora & Ottmann, 2020a:171). Anche questi rischi, per essere ridotti al minimo, necessitano della definizione di una gestione del

rischio sistematica (ibidem). Un rischio sia per il cliente (o per il destinatario della traduzione) che per il traduttore stesso è collegato alla commissione di errori nella traduzione: se presenti nel lavoro consegnato al committente, possono portare a una perdita di reputazione, conseguenze legali, danni a persone o cose (Canfora & Ottmann, 2018, in Canfora & Ottmann, 2020a:171-172). Tuttavia, non è ancora stato chiarito in modo definitivo chi sia responsabile in caso di danno causato dall'uso della MT (Canfora & Ottmann, 2020a:171). È possibile identificare due diversi principi di responsabilità: responsabilità con colpa e responsabilità senza colpa (ivi, 173). Nel primo caso la responsabilità si basa sul comportamento illecito dell'uomo, quindi, sull'errore umano che deve essere dimostrato (ibidem). Nel secondo caso la responsabilità, chiamata "responsabilità oggettiva" ricade su un operatore che risponde, indipendentemente da qualsiasi errore umano, di tutti i danni comparsi durante il processo di traduzione: per questo motivo dovrebbe essere sempre presente un'assicurazione a copertura della responsabilità civile (ibidem).

La traduzione, dunque, è un processo che richiede costante attenzione: errori che potrebbero sembrare banali possono rivelarsi fatali. Un maggior riguardo deve essere prestato nei confronti dei dati inseriti online

### **2.6.2 Protezione dei dati**

Una grande problematica all'interno del post-editing della traduzione automatica è rappresentata dalla protezione dei dati personali.

Kamocki e Stauch (2020:43) si propongono di fare chiarezza su questa tematica, fornendo, innanzitutto, la definizione di "dato personale". Essi fanno riferimento al GDPR "General Data Protection Regulation" (EU Regulation 2016/679), che dal 2018 è la legislazione europea chiave applicabile al trattamento dei dati personali in modo generale. All'articolo 4(1) è indicato che con la dicitura "dato personale" si intende "any information relating to an identified or identifiable natural person"<sup>12</sup> (ibidem). La norma include i dettagli personali (come il proprio nome e cognome, data di nascita, ecc.), informazioni di contatto (ad esempio, e-mail, indirizzo, numero di telefono, ecc.), e qualsiasi altra informazione relativa a persona vivente (ivi, 43-44). Il GDPR si applica indifferentemente che il trattamento venga messo in pratica per ragioni di pubblico

---

<sup>12</sup> "Qualsiasi informazione relativa a una persona naturale indenticata o identificabile" (tradotto da me).

interesse, per ragioni commerciali, o addirittura, in casi più rari, per spiare le persone (ibidem). Come si può facilmente comprendere, il trattamento dei dati personali è da applicarsi anche quando si parla di traduzione automatica, in tutte le sue fasi (ivi, 44-45):

- nella fase di sviluppo, in quanto è altamente probabile che i corpora paralleli, presi da fonti tra loro diverse e utilizzati per addestrare i motori di traduzione automatica, contengano dati personali (ibidem);
- nella fase di *deployment*, quando i dati personali potrebbero essere presenti all'interno del sistema di traduzione automatico o procurati dall'utilizzatore (ibidem);
- nella fase di valutazione, quando i dati presenti sono usati per valutare qualitativamente e quantitativamente le prestazioni del sistema di traduzione automatica (ibidem);
- potenzialmente anche nella fase di *repurposing*, in cui i dati inseriti dall'utente possono essere utilizzati per uno scopo non collegato alla traduzione automatica (ad esempio, per la profilazione o il marketing diretto) (ibidem).

Non tutti i sistemi di traduzione automatica proteggono i testi e i dati che si vanno ad inserire: per quanto riguarda i sistemi *in-house* (interno all'agenzia di traduzione, per esempio), questi risultano abbastanza protetti, poiché i testi inseriti nel sistema sono conservati in modo sicuro in un ulteriore sistema interno o in un server; oppure, ancora, i sistemi *cloud-based* risultano essere anch'essi affidabili, in quanto ci si accede solo tramite un codice di sicurezza (Nitzke & Hansen-Schirra, 2021:54). Al contrario, è da prestare maggiore attenzione ai sistemi esterni, che non sono dotati di un codice di sicurezza, o ai sistemi di traduzione automatica online e gratuiti (un esempio noto a chiunque è rappresentato da Google Translate), grazie ai quali la MT è diventata accessibile a chiunque (ibidem). Inevitabilmente, quando si traduce un testo in uno di questi strumenti online, si inseriscono dei dati, che possono essere delle tipologie più disparate: contenuti di social media, articoli di giornale, corrispondenze private e professionali. All'interno di essi sono compresi i dati sensibili e personali dell'utente, che, se non adeguatamente protetti tramite un sistema di sicurezza, sono facilmente accessibile a terzi (Kamocki *et al*, 2015:270). Dunque, i sistemi di traduzione

automatica che non permettono di lavorare in un ambiente tutelato, devono essere evitati (ibidem).

È sempre necessario chiedere l'autorizzazione del cliente prima di procedere alla traduzione automatica di un determinato testo, specificatamente se si vuole utilizzare un sistema di traduzione automatica online:

Private users should consider translating only those bits of texts that do not contain any information relating to third parties (which in practice may limit them to translating text into, and not from, a different language to their own). Businesses in particular may find such a limitation rather constricting and to protect their own data and the data of their clients, may prefer to opt for a payable offline MT tool instead of a 'free' online service<sup>13</sup>. (Kamocki et al., 2015:279)

Secondo Kamocki *et al.* (2015:279), dunque, se si usano sistemi di traduzione automatica online è consigliato evitare l'inserimento dei dati personali del cliente: in alternativa è consigliabile utilizzare un servizio offline, seppure a pagamento, ma che possa offrire una sicurezza maggiore tanto per il cliente quanto per il traduttore.

I sistemi di traduzione automatica forniscono, nei loro termini di servizio, informazioni dettagliate relativamente alla sicurezza dei dati. Nitzke e Hansen-Schirra (2021:54-55) riportano le direttive indicate sul sito della compagnia tedesca DeepL, che fornisce sistemi di traduzione automatica sia gratuitamente che a pagamento. A seconda della gratuità o meno del servizio, i termini presentano delle differenze: per quanto riguarda la versione gratuita, è richiesto all'utente di essere certo di poter inserire all'interno del server i testi di cui si desidera ottenere la traduzione, poiché vengono usati per implementare la qualità delle traduzioni prodotte da DeepL; viene ribadito che non devono essere inseriti testi contenenti dati personali (DeepL.com). La versione "pro", a pagamento, assicura all'utente una maggiore sicurezza nella protezione dei dati personali, se inseriti: nella versione a pagamento, infatti, i testi e dati inseriti non vengono utilizzati per implementare la qualità dell'output (ibidem).

## **2.7 Prospettive future**

Nell'attuale periodo storico, come precedentemente sottolineato, si assiste a una continua evoluzione e miglioramento nel campo della traduzione automatica e del post-

---

<sup>13</sup> Gli utenti privati dovrebbero prendere in considerazione la possibilità di tradurre solo quelle porzioni di testo che non contengono informazioni relative a terzi (che, in pratica, potrebbe limitarli a tradurre testi in, e non da, una lingua diversa dalla propria). Le aziende, in particolare, potrebbero trovare tale limitazione piuttosto restrittiva, e, per proteggere i propri dati e quelli dei propri clienti, potrebbero optare per uno strumento di traduzione automatica offline a pagamento, anziché utilizzare un servizio online "gratuito" (tradotto da me).

editing, grazie a un sempre più rapido progresso tecnologico dell'intelligenza artificiale. Non risulta difficile immaginare, difatti, che la qualità degli output prodotti dai sistemi di traduzione automatica, in un futuro non troppo lontano, saranno sempre più accurati e la traduzione ancora più attinente al testo originale e al suo significato.

Alcuni traduttori, di fronte a questo veloce sviluppo, si sentono minacciati da tale avanzamento tecnologico, arrivando a temere che presto le macchine si sostituiranno ai traduttori umani, ponendo fine, ovvero, a una professione esistente da migliaia di anni. Esiste, tuttavia, una reale possibilità che ciò avvenga, o può questa essere semplicemente considerata come una visione utopistica delle future prospettive lavorative in ambito traduttivo? Il consenso generale afferma che l'intelligenza artificiale avrà un grosso impatto (e lo sta già avendo, in realtà) su diverse tipologie di impieghi lavorativi, tra cui, ovviamente, il mestiere del traduttore (O'Brien, 2021:182). La posizione di quest'ultimo, ciononostante, non sarà da considerarsi come superflua; tuttavia, subirà (e sta attualmente subendo) un cambiamento nei compiti da svolgere (ibidem). La stessa figura del post-editor, in effetti, può essere interpretata come un'evoluzione del ruolo del traduttore, occupata nello svolgimento di una serie di compiti essenziali: i sistemi di traduzione automatica, in genere, non sono ancora in grado di produrre un output perfetto, privo di errori, e, secondo alcuni esperti, non è da aspettarsi che questo livello di correttezza dell'output venga raggiunto in tempi brevi (ibidem). Quella del traduttore continuerà ad essere una figura fondamentale, e numerosi contenuti vengono attualmente tradotti senza l'uso di un sistema di traduzione automatica: infatti, la traduzione automatica non sarà in grado di fornire contenuti creativi, capace di "parlare" al possibile utente finale (ibidem). Ciò non indica che la MT non possa essere utilizzata per contenuti creativi; piuttosto significa che i suggerimenti proposti dal sistema MT potranno fungere da base, come un "trampolino di lancio" per creare qualcosa di più creativo del suggerimento stesso, in cui l'intervento umano rappresenterà ancora una parte cruciale per diversi anni a venire (ibidem).

La traduzione costituisce una delle più grandi sfide per l'intelligenza artificiale fin dai suoi albori. Continua a essere una grossa complessità l'esecuzione di una traduzione, effettuata da una macchina, che sia, però, comparabile a quella compiuta da un traduttore umano: "that's because language is the most natural thing for humans"<sup>14</sup>,

---

<sup>14</sup> "La lingua è la cosa più naturale per gli umani" (tradotto da me).

afferma Marco Trombetti, il CEO dell'azienda "Translated", ponendo particolare attenzione alle difficoltà esistenti nel riprodurre la lingua naturale attraverso l'uso dell'intelligenza artificiale (Translated.com). Tuttavia, Marco Trombetti afferma con sicurezza che non si è lontani dal raggiungimento della *singularity*, ovvero ci si sta avvicinando alla chiusura del divario tra quello che può produrre un traduttore esperto, umano, e ciò che può essere generato da un sistema di traduzione automatica (ibidem). Translated è un'azienda specializzata in traduzione e che, in particolare, combina l'intelligenza artificiale e la capacità umana per ottenere dei risultati il più possibile vicini alla perfezione. L'azienda, nel corso degli anni, ha costantemente monitorato e valutato la qualità delle traduzioni prodotte dai sistemi di traduzione automatica, arrivando, nel 2011, a stabilire una metrica chiamata "Time to Edit" (TTE), ovvero il tempo richiesto dai traduttori professionisti più performanti al mondo per controllare e correggere le traduzioni suggerite dalla traduzione automatica (ibidem). Il TTE viene considerato come il miglior modo per misurare la qualità della traduzione, in quanto fornisce un'indicazione di quale sia il tempo medio reale necessario al controllo e alla correzione di una traduzione (ibidem). Il TTE rappresenta una svolta come indicatore della qualità della traduzione rispetto a BLEU (Bilingual Evaluation Understudy) o COMET (Crosslingual Optimized Metric for Evaluation of Translation), indicatori precedentemente utilizzati per il calcolo dello sforzo cognitivo richiesto durante la correzione di una traduzione (ibidem). Translated prevede che il TTE, se la sua tendenza continuerà a diminuire con lo stesso ritmo del 2014, scenderà a un secondo nei prossimi anni, avvicinandosi a quella che potrebbe essere definita come una traduzione "perfetta" (sebbene sia prodotta da un sistema di traduzione automatica), la quale si ottiene, dunque quando il tempo impiegato da un traduttore professionista per il post-editing di una traduzione automatica è uguale a quello che impiegherebbe per la revisione di un testo prodotto da un collega, ovvero da un traduttore umano (ibidem). Il TTE, tuttavia, è influenzato da altri due fattori, oltre alla qualità dell'output della traduzione assistita, ossia dall'evoluzione dello strumento di editing e dalla qualità proposta dai traduttori (ibidem).

Translated, per la formulazione di questa ipotesi sul futuro della traduzione automatica e del post-editing, ha raccolto nel corso degli ultimi anni oltre 2 miliardi di modifiche su frasi effettivamente tradotte in un contesto di produzione, modifiche

apportate da 136.000 freelance, tra i più bravi al mondo, che hanno collaborato con Translated, utilizzando il loro strumento di traduzione assistita (CAT), Matecat (ibidem).

Translated, grazie ai dati raccolti e alla collaborazione costante tra intelligenza artificiale e traduttori umani professionisti, esistente da oltre 20 anni, è in grado di fornire una visione futura sul ruolo del traduttore e sull'evoluzione dei suoi compiti: l'azienda assicura con certezza che il raggiungimento della *singularity* porterà enormi vantaggi, incrementando la domanda di MT di 100 volte, comportando, conseguentemente, un aumento di richiesta di traduttori altamente specializzati per il controllo della qualità dell'output prodotto e il post-editing di quest'ultimo (ibidem).

L'estrema globalizzazione del mondo di oggi implica una sempre maggiore domanda di traduzione dei contenuti in più lingue possibili, affinché siano accessibili a un pubblico vasto. La figura del traduttore, dunque, continua ad avere una fondamentale importanza e continuerà ad averla nel futuro: probabilmente, sussisterà sempre la necessità di un revisore umano dell'output della traduzione prodotto da un sistema MT, in quanto difficilmente si presenterà con una qualità pari a quella che potrebbe essere fornita da un traduttore professionista, sebbene lo possa sembrare. Per tale ragione, il ruolo del post-editor sta acquisendo una posizione vantaggiosa, essendo adeguata all'attuale evoluzione tecnologica in atto.

## **2.8 Fattori di (non) adozione del post-editing**

Oggi giorno i sistemi di traduzione automatica neurale producono un output di buona qualità, come precedentemente e ripetutamente dichiarato. Ciononostante, non è ancora possibile affermare che tale output possa essere utilizzato senza prima essere rivisto, e, dunque, senza essere sottoposto a post-editing. Il mercato della traduzione, in seguito a un maggiore utilizzo dei sistemi di traduzione automatica, sta riscontrando un forte aumento della domanda di post-editing, una pratica divenuta estremamente richiesta e diffusa, in particolare nel corso degli ultimi anni. Numerosi traduttori, pur ricevendo diverse richieste di post-editing della traduzione automatica, si dimostrano tuttora ostili nei confronti di questo lavoro, ritenendolo una perdita di tempo e di denaro. In modo particolare, viene spesso messa in discussione l'utilità della *machine translation* stessa.

Secondo uno studio di Cadwell *et al.* (2017:310), alcuni traduttori affermano che il post-editing, con certe tipologie di testi, consente un aumento di velocità e produttività, risultando particolarmente efficace con determinate coppie linguistiche, producendo un

output di buona qualità; diversi traduttori, inoltre, asseriscono che la traduzione automatica può, in alcuni casi, essere fonte di ispirazione, fornendo nuove idee, assistenza terminologica, e aiutando a comprendere delle lingue *source* che non si padroneggiano particolarmente bene (ibidem); altri, infine, dichiarano di voler contribuire a un miglioramento della qualità dell'output della traduzione automatica e per questo ne fanno utilizzo (ibidem). Nello stesso studio (di Cadwell *et al.*, 2017:310) viene anche evidenziata la necessità di occuparsi della traduzione automatica, e perciò anche del post-editing, in quanto la sua adozione, al momento attuale, risulta inevitabile e, in alcuni circostanze, addirittura obbligatoria. Ciononostante, sono numerosi i traduttori ad aver riportato delle motivazioni a sfavore dell'uso e dell'adozione della traduzione automatica: il suo utilizzo appare inadatto con alcuni tipi di testi, e, con determinate coppie linguistiche, l'output prodotto risulta di scarsa qualità (ivi, 311). Inoltre, diversi traduttori sostengono che un eccessivo utilizzo della traduzione automatica induce il traduttore a commettere più errori, anziché meno, influenzando negativamente le sue abilità (ibidem). Si ritiene che il post-editing, rispetto alla traduzione o alla revisione, necessiti di maggiore concentrazione, e, in alcuni casi, allunghi i tempi di revisione e modifica: infatti, non è raro che i traduttori ripongano poca fiducia nei sistemi di traduzione automatica e nella qualità dell'output prodotto (ibidem). Il post-editing, dunque, è rappresentato come un'attività poco piacevole, che, oltretutto, svaluta il lavoro del traduttore e diminuisce la creatività del traduttore stesso (ibidem). Una componente che rende tale pratica poco apprezzata è la paura che i sistemi di traduzione automatica possano completamente sostituire i traduttori umani, data la mancanza di una grande interazione umana nel testo, e ciò spinge alcuni traduttori a non volersi servire di tale tecnologia (ibidem). Viene comunque sottolineato che la MT potrebbe fornire delle idee sulla traduzione stessa o essere utilizzata come punto di partenza nel processo di traduzione (ivi, 312).

Uno dei principali motivi per cui numerosi traduttori non si vogliono occupare di post-editing è la mancanza di un metodo di pagamento adeguato allo sforzo che viene richiesto per effettuare il post-editing stesso, sebbene non sia stato menzionato dalla maggioranza dei partecipanti allo studio di Cadwell *et al.* (2017:312): tuttavia, gli autori affermano che i partecipanti allo studio erano lavoratori dipendenti ed è possibile che le questioni di sicurezza e retribuzione siano più significative in altri contesti istituzionali

o commerciali (ad esempio, traduttori freelance che lavorano per un fornitore di servizi linguistici o governativi - LSP - o per un'istituzione governativa) (ibidem).

In uno studio meno recente, condotto da Guerberof Arenas nel 2013, la ricercatrice chiese ad alcuni traduttori quale fosse la propria opinione sulla MT e sul PE. La maggior parte dei partecipanti affermò di aver notato, con il passare del tempo e maturando esperienza nel campo, un aumento della produttività nel post-editing, o perlomeno una produttività costante: tali dichiarazioni non stupiscono, in quanto il post-editing è un'attività fortemente ripetitiva e, grazie a un continuo allenamento, in determinati casi risulta più semplice e più veloce rintracciare gli errori presenti nel testo, spesso ricercando gli stessi schemi precedentemente affrontati, come affermato dalla maggior parte dei partecipanti (Guerberof Arenas, 2013:77-78). La ricercatrice, inoltre, si volle soffermare su un altro punto, ovvero lo sforzo richiesto ai post-editor nel loro lavoro di modifica e revisione: circa il 40% degli intervistati dichiarò che lo sforzo richiesto dal post-editing fosse superiore a quello necessario per effettuare una revisione (ivi, 78); tuttavia, una parte degli intervistati, pari al 20% dei partecipanti, sostenne lo sforzo cognitivo richiesto da questa tipologia di lavoro fosse minore (paragonato allo sforzo cognitivo richiesto dalla revisione) (ibidem). Infine, ai partecipanti dello studio fu chiesto quale fosse il proprio grado di soddisfazione del pagamento ricevuto per il post-editing: 9 traduttori tra i 24 partecipanti alla ricerca dissero di ritenersi insoddisfatti o molto insoddisfatti dei guadagni ottenuti (ivi, 79). È importante ricordare che tale indagine fu effettuata ben dieci anni fa, e all'epoca l'unico metodo di pagamento proposto, perlomeno da ciò che si evince dalle domande e risposte di tale questionario, era il pagamento tramite il calcolo del prezzo a parola. Oggigiorno, invece, vi sono diverse metodologie di remunerazione del post-editing, ed è tuttora in corso un acceso dibattito su quale tipologia di pagamento sia la più equa e adatta. Venne, poi, riportato da Guerberof Arenas (2013:79) che alcuni traduttori fornirono delle spiegazioni in merito alla loro insoddisfazione sui guadagni ottenuti: ad esempio, un post-editor dichiarò che, a causa della recessione economica dell'epoca, i prezzi erano scesi e guadagnava molto meno di quando aveva iniziato; un'altra traduttrice affermò che il guadagno dipendeva dall'incarico, dal cliente o dall'agenzia, in quanto ognuno di questi offriva un servizio di pagamento differente (ibidem).

Anche in un'ulteriore ricerca, svolta da Sakamoto nel 2019, molti dei partecipanti allo studio hanno espresso pareri prevalentemente negativi sul post-editing della *machine translation* (Sakamoto, 2019:205-206). Tra i partecipanti vi erano vari project manager che, infatti, si dichiararono contrari all'offrire servizi di MTPE, e solamente un partecipante riconobbe nel post-editing il destino futuro della traduzione (ivi, 206). I project manager partecipanti affermarono di avere il timore che i loro traduttori freelance non sarebbero stati disponibili a svolgere il lavoro di post-editing principalmente per due motivi: il PE è infatti generalmente percepito come un lavoro noioso e, secondo alcuni, addirittura "avvilente" (Moorkens and O'Brien 2017:109, in Sakamoto, 2019:206); inoltre, il post-editing comporta una retribuzione inferiore rispetto alla traduzione tradizionale (umana) (Sakamoto, 2019:206). Tuttavia, i project manager partecipanti affermarono di essere coscienti dell'aumento della richiesta di servizi di MTPE da parte dei clienti, sempre più esperti nel campo della tecnologia, al fine di ottenere una maggiore produttività e un abbassamento dei costi, attraverso tariffe più basse (ibidem). Ciò, però, porta a un inevitabile abbassamento dei pagamenti per il lavoro di post-editing: anche in questo caso i partecipanti espressero la loro paura nei confronti della MT, ovvero quella di perdere i propri traduttori proprio a causa di un pagamento troppo scarso (ibidem).

Infine, un grande problema che vale la pena citare e che riguarda diversi post-editor è l'*over-editing*, termine con cui si intende il fenomeno dei cambiamenti preferenziali, sia nel post-editing che nella revisione (Nitzke & Gros, 2021:21). Si parla di *over-editing* quando i traduttori rivedono o effettuano il post-editing di un output più del necessario, sebbene vengano date alcune linee guida da seguire (ibidem). Sia nel caso della revisione che in quello del post-editing, molti traduttori sentono la necessità di modificare tutto ciò che, secondo la loro opinione, non risulta corretto, migliorando gli aspetti linguistici al fine di raggiungere una qualità perfetta: tuttavia, dovrebbero attenersi al livello di post-editing richiesto e tralasciare determinati errori presenti, se non diversamente specificato (ibidem). È necessario, infatti, seguire le linee guida affinché il processo di post-editing (o anche quello di revisione), sia efficiente e utile, soprattutto in termini di compensazione: il traduttore, infatti, non verrà pagato per le modifiche aggiuntive e non richieste, che vengono apportate (ibidem). Il fenomeno dell'*over-editing* è da anni riconosciuto specialmente nei sistemi CAT, in particolare

nelle TM, in quanto i traduttori frequentemente modificano i *match* proposti, anche nel caso in cui si tratti di *exact match*, rivedendo fin troppo precisamente i *fuzzy match*, sebbene in molti casi necessitino solamente di piccoli miglioramenti (ibidem). Molto spesso, come dimostrato in alcuni studi, si tratta di modifiche preferenziali a livello sintattico o lessicale, poiché è “normale” che un traduttore, trovandosi di fronte a una frase non tradotta da sé stesso, veda errori anche quando questi non sussistono (ibidem). Dunque, è necessario attenersi alle indicazioni previste per il *light* e il *full* PE, al fine di non perdere tempo prezioso e, conseguentemente, denaro, sebbene la qualità dell’output risulti inferiore a quella che si vorrebbe personalmente ottenere (ivi, 31). Non di rado, i traduttori, nel confrontare il testo sorgente con l’output prodotto dal sistema di traduzione automatica, si autopropongono una versione della traduzione diversa: per tale ragione risulta automatico voler modificarla secondo le proprie idee (ibidem). Una corretta formazione nell’ambito del PE permette di allenare i traduttori, affinché possano evitare l’*over-editing* e raggiungere, quindi, l’obiettivo richiesto con un giusto rapporto tra sforzo impiegato e il guadagno previsto (ivi, 32).

### **III. Indagine sulla compensazione equa nel post-editing**

#### **3.1 Introduzione**

Il pagamento del post-editing è un argomento molto dibattuto nell'industria della traduzione, sul quale, purtroppo, non è ancora stata fatta chiarezza: non vi è, infatti, un'unica metodologia per la retribuzione di questo lavoro e le decisioni in merito ricadono spesso nelle mani dei clienti, non di rado mal informati su tale ambito. Il post-editing, infatti, viene visto frequentemente come un'opportunità per i clienti di poter risparmiare sui costi, senza tenere conto dell'effettiva complessità di questo processo e di tutte le fasi che lo compongono. Esistono svariate opinioni sul tema e l'obiettivo che il presente lavoro di tesi si ripropone è proprio quello di raccogliere i vari punti di vista e analizzarli. Lo studio in questione ha un duplice scopo: da un lato esaminare la percezione del post-editing e della sua remunerazione da parte del pubblico generale, che vi ha a che fare in modo indiretto e tra cui potrebbero esserci dei potenziali clienti; dall'altro lato, invece, è volto a interagire direttamente con i post-editor raccogliendo, innanzitutto, le loro opinioni sul post-editing stesso (quindi anche in questo caso si vuole esaminare la percezione che i post-editor hanno del proprio lavoro), e, in secondo luogo, comprendere il loro punto di vista sui diversi metodi di retribuzione attualmente esistenti, esprimendo la propria opinione su quello che ritengono maggiormente equo e adatto. Il questionario si basa sull'iniziale presupposto che vi sia una scorretta percezione del post-editing e della sua remunerazione da parte del pubblico generale e un'insoddisfazione complessiva dei post-editor per quanto riguarda i mezzi di pagamento utilizzati e secondo i quali vengono remunerati. Si ritiene, dunque, che il lavoro di post-editing non sia adeguatamente pagato, e che, di conseguenza, i post-editor non guadagnino abbastanza. Un altro presupposto di partenza, inoltre, è che la maggioranza dei post-editor venga remunerata secondo un pagamento a parola, il quale, tuttavia, è spesso giudicato inadeguato a questa tipologia di lavoro. Obiettivo dello studio, dunque, è anche comprendere se l'adozione generale di una remunerazione su base oraria potrebbe essere un metodo di pagamento che accolga le richieste ed esigenze della maggior parte dei post-editor e, quindi, se questo metodo di remunerazione potrebbe essere il migliore e divenire il prescelto per tale tipologia di lavoro.

Il questionario è stato sottoposto a un totale di 132 partecipanti, divisi omogeneamente in due gruppi da 66 persone ciascuno: uno composto dai post-editor e

l'altro dal pubblico generale. Dunque, lo studio comprende anche non esperti del settore, con una formazione educativa e lavorativa tra loro stessi differente. Non è stato previsto un limite di età: per questo i partecipanti rappresentano una fascia d'età che va dai 18 agli oltre 65 anni (divisi in sei diversi gruppi: dai 18 ai 24 anni, dai 25 ai 34 anni, dai 35 ai 44 anni, dai 45 ai 54 anni, dai 55 ai 64 anni e gli over 65). Non vi sono state limitazioni di genere e di titolo di studio, ma tali dati sono stati richiesti per sottolineare eventuali differenze di concezione tra i partecipanti o per fornire un eventuale valore aggiuntivo alle affermazioni riportate. Non è stata imposta alcuna restrizione dal punto di vista della provenienza geografica: infatti, il questionario è stato redatto in lingua inglese proprio per venire incontro alle esigenze linguistiche dei diversi partecipanti. Sono state contattate numerose associazioni di traduttori e interpreti, con sede in diverse parti del mondo, ottenendo la partecipazione attiva di alcuni dei loro membri. Le associazioni in questione sono: Asociación Panameña de Traductores e Intérpretes (Panama), New Zealand Society of Translators and Interpreters (Nuova Zelanda), Associazione Svizzera Traduttori, Terminologi e Interpreti, Bundesverband der Dolmetscher und Übersetzer e.V. (BDÜ) (Germania), Unione Ceca dei Traduttori e Interpreti (Repubblica Ceca), Estonian Association of Masters in Conference Interpreting and Translation (Estonia), Israel Translators Association (Israele), Association of Scientific and Technical Translators of Slovenia, oltre all'associazione italiana AITI (Associazione Italiana Traduttori e Interpreti). L'obiettivo, infatti, non è stato quello di focalizzare l'attenzione su una determinata area geografica, ma, piuttosto, quello di raccogliere il numero maggiore di percezioni e idee dei singoli partecipanti sul tema. Per quanto riguarda il pubblico generale, è stato proposto il questionario a conoscenti, amici e parenti, pubblicandolo, inoltre, su alcuni canali social (ad esempio, WhatsApp e Instagram).

Il questionario può essere suddiviso principalmente in due parti: la prima, sottoposta sia ai post-editor che al pubblico generale, è dedicata alla comprensione della percezione sul post-editing, all'analisi della conoscenza di questo lavoro e dei metodi di remunerazione esistenti; la seconda parte, invece, è stata compilata solamente dai post-editor, in quanto è volta a esaminare le loro opinioni personali sui metodi di pagamento, al fine di comprendere quale possa essere il sistema di retribuzione più equo per il lavoro di post-editing. Una piccola ultima sezione è stata dedicata ai dati personali (età,

genere e titolo di studio), offrendo la possibilità ulteriore ai partecipanti di esprimere alcune opinioni individuali sulla tematica. La prima parte è stata progettata sulla base del modello dell'associazione libera (vedi §3.3 Metodologia: associazione libera), mentre la seconda parte rispecchia un tipico questionario basato sull'analisi quantitativa. Prima di inoltrarsi nel vivo dell'analisi delle risposte ottenute, viene proposta un'introduzione ai diversi metodi di pagamento utilizzati nel post-editing, e, in seguito, cosa si intende per "associazione libera" e il suo funzionamento.

### **3.2 Metodi di pagamento del post-editing**

Uno degli argomenti più dibattuti nell'ambito del post-editing, in modo particolare nel corso degli ultimi anni, è la compensazione di tale lavoro. Il dibattito iniziò, come già precedentemente affrontato, con i sistemi CAT per l'adattamento del sistema di pagamento in base ai diversi tipi di *match* ottenuti (O'Brien, 2021:179-180). In modo simile, al giorno d'oggi tale questione viene affrontata nell'ambito del post-editing della MT (ivi, 180). Inizialmente, molti dei clienti ritennero che il pagamento a parola potesse rappresentare il miglior metodo di remunerazione per questo lavoro (ibidem). Tuttavia, dopo un iniziale adeguamento a questo metodo di pagamento, molti traduttori si opposero, in quanto, in alcuni casi, si dimostrava inadatto, poiché la produttività nel post-editing dipende da una serie di fattori, quali la coppia linguistica, il sistema di traduzione automatica utilizzato, il tipo di contenuto e così via (ibidem). Al momento attuale, tale discussione non si può ancora ritenere risolta, sebbene siano numerosi i traduttori che concordano che il miglior metodo di remunerazione sia quello su base oraria (ibidem).

Essendo il post-editing molto diverso dalla traduzione manuale e, in parte, anche dalla traduzione assistita, vi è la necessità di definire un sistema di pagamento adeguato al lavoro richiesto, che possa rispettare le esigenze dei post-editor. Attualmente esistono in particolare tre metodi di pagamento: il pagamento a parola, che può essere considerato come uno dei metodi più classici, utilizzato soprattutto nella traduzione, il pagamento basato su una tariffa oraria e, infine, il pagamento basato sulla distanza di modifica, meglio conosciuta con la sua denominazione inglese *edit-distance*.

Ognuno dei singoli modelli di pagamento presenta una serie di vantaggi e svantaggi e, anche a causa di tale motivazione, risulta tuttora difficile stabilire quale sia il metodo di remunerazione più adatto. Spesso, è proprio a causa della retribuzione scarsa e non

adeguata, se confrontata con lo sforzo impiegato per lo svolgimento del lavoro, che molti traduttori si rifiutano di accettare incarichi di post-editing. Tuttavia, vi è la certezza che, all'interno del mercato della traduzione, vi sarà una sempre maggiore richiesta di post-editing (Grizzo, 2020:367): infatti, quella del post-editing della MT è stata l'area nel mercato della traduzione con la crescita più rapida dal 2010 in poi, prima ancora del passaggio dalla traduzione automatica statistica a quella neurale (Moorkens, 2022:132). Spesso sono i clienti stessi a stabilire la metodologia di compensazione e il post-editor si trova, dunque, davanti a un'impossibilità di scelta. Un post-editor esperto, d'altra parte, è anche in grado di capire quale metodo di remunerazione sia più adeguato al suo lavoro e comprendere come vorrebbe essere retribuito (Grizzo, 2020:367). Vi sono alcuni traduttori che hanno una visione totalmente positiva del post-editing, ritenendolo un lavoro utile e redditizio; altri, invece, lo trovano avvilente, sostenendo che il post-editing svaluta il lavoro del traduttore. Inoltre, sostengono che normalmente i metodi di pagamento utilizzati non sono soddisfacenti, arrivando a giudicare tale impiego una perdita di tempo e denaro, non trovando adeguato riconoscimento e senso di realizzazione nel suo svolgimento (Moorkens, 2022:133). Il punto di vista dei post-editor/traduttori insoddisfatti è da tenere altamente in considerazione, poiché, nel caso in cui ciò non avvenga, si potrebbe incorrere in una carenza non solo di post-editor, ma anche di traduttori stessi: è fondamentale soddisfare le esigenze, se possibile, di tutti i soggetti interessati, in modo particolare perché si tratta di una grossa porzione di traduttori (ibidem). Eppure, non è raro che i clienti mettano al primo posto la produttività e la riduzione dei costi, anziché la soddisfazione dei post-editor che lavorano per loro (ibidem).

Normalmente, ogni azienda decide un proprio metodo di pagamento per il post-editing. Un esempio precedentemente affrontato è quello dell'azienda "Translated", il cui CEO Marco Trombetti in un'intervista ha dichiarato che i post-editor che lavorano con Translated vengono remunerati in base al numero di parole, ma gli sconti percentuali sulle parole sono calcolati in base al tempo effettivo speso per il post-editing (Slator.com). Nel corso dell'intervista ha spiegato, inoltre, che il riconoscimento del valore dei traduttori e post-editor, e del loro contributo, sono da sempre fondamentali per l'azienda, aggiungendo che fin dall'inizio dell'utilizzo della traduzione automatica, i traduttori freelance in Translated sono sempre stati pagati sia per le parole tradotte da

loro che per quelle processate dal sistema di traduzione automatica, portando a un aumento del compenso del 25% all'ora nei confronti dei post-editor (ibidem). Ciò è dovuto al fatto che i traduttori, senza MT, impiegherebbero circa il 50% in più del tempo che impiegano per il post-editing; dunque, il post-editor non viene pagato per il 50% di parole in meno, ma ne scontano solo il 25% (ibidem). Pur mantenendo una retribuzione a parola e non su base oraria, che, come suggerito dall'intervistatore, sembrerebbe essere la soluzione migliore nel settore del post-editing, l'azienda "Translated" sembrerebbe venire incontro alle esigenze dei traduttori e assicurare loro una compensazione equa (ibidem).

Tuttavia, come visto precedentemente, non è raro che il lavoro di post-editing non risulti, in genere, abbastanza soddisfacente: il metodo di pagamento, infatti, spesso non tiene conto delle diverse problematiche si incontrano durante il post-editing (ovvero la scarsa qualità dell'output o il tempo necessario a produrre un lavoro di buona qualità), portando, in alcuni casi, a guadagni nettamente inferiori rispetto a quelli normalmente ottenuti con la traduzione o con la revisione. Vengono di seguito analizzate le tre metodologie di remunerazione più presenti sul mercato del post-editing, evidenziando, per ognuna di esse, i punti a favore e quelli a sfavore.

### **3.2.1 Pagamento a parola**

Il pagamento a parola, utilizzato nel lavoro di post-editing, si basa, in realtà, sul metodo di pagamento tradizionale degli strumenti CAT (Grizzo, 2020:363). Negli strumenti CAT, infatti, durante la preparazione di un progetto, tutti i segmenti il cui *match* risulta al di sotto di una certa soglia (ad esempio, 75%) e quelli non presenti all'interno di una memoria di traduzione vengono pre-tradotti tramite una MT, non rimanendo, dunque, vuoti (ibidem). Il sistema registra la quantità (il numero) di nuove parole inserite, prodotte dal sistema di traduzione automatica (ibidem). Questo consente di poter capire quante parole debbano essere effettivamente remunerate: normalmente il prezzo per parola si aggira tra il 40% e il 60% del prezzo normalmente utilizzato (nella traduzione) per una nuova parola, e solitamente viene negoziato anticipatamente (ivi, 364). Tuttavia, il prezzo viene adattato a una serie di variabili: viene infatti deciso in base al livello di post-editing applicato (ovvero *light* o *full* post-editing), alla coppia linguistica utilizzata e al contenuto trattato (ibidem). Dunque, è necessario effettuare, prima di tutto,

un'analisi, al fine di valutare i criteri sopracitati per calcolare rapidamente l'importo totale del nuovo progetto, così da fissarlo in anticipo (ibidem).

Il pagamento a parola è un metodo molto vantaggioso per il cliente, in quanto gli fornisce una certezza riguardo ai costi dell'incarico (ibidem). Tuttavia, per il post-editor potrebbero presentarsi una serie di svantaggi: infatti, se l'output prodotto dal sistema di traduzione automatica presenta una qualità scadente, il post-editor corre il rischio di lavorare in modo antieconomico, ovvero incorrere in un testo nel quale è necessario applicare un numero troppo elevato di modifiche, dovendo addirittura riscrivere traduzioni di interi segmenti, spreco, in questo modo, molto tempo senza venire tuttavia retribuito (ibidem). Ciononostante, nel testo a cui si fa riferimento in questo paragrafo, l'autrice Sara Grizzo afferma in modo positivo di aver riscontrato, nel corso della sua carriera, che in ogni lavoro di post-editing vi sono dei segmenti che possono essere considerati buoni e altri che lo sono meno; aggiunge, inoltre, che se la qualità dell'output della MT risulta molto scadente probabilmente significa che quel testo non è adatto a essere tradotto con un sistema MT o non è stato usato il sistema giusto (ibidem).

### **3.2.2 Pagamento su base oraria**

Il pagamento su base oraria è molto probabilmente il modello di retribuzione più apprezzato dalla maggior parte dei post-editor (Grizzo, 2020:364). Il pagamento su base oraria può avvenire in due modi, in base al calcolo della durata del lavoro: infatti, la prima opzione prevede l'annotazione delle ore di lavoro da parte del post-editor stesso, che le comunicherà al cliente al momento della consegna del lavoro svolto; la seconda alternativa prevede la richiesta da parte del cliente dell'uso di uno strumento di registrazione delle ore (ibidem). Grizzo (2020:364) riporta che quest'ultima funzione è già integrata all'interno di alcuni strumenti CAT, attraverso la quale, quando il progetto risulta completato, può essere generato un report che mostra le ore di lavoro svolte e la produttività oraria: ad esempio, in MemoQ, oltre alle ore di lavoro svolte, vengono indicati il numero di parole e di caratteri di cui si è effettuato il post-editing (ibidem). Tale approccio, pur fornendo una maggiore sicurezza al cliente, non viene apprezzato da molti post-editor, che si sentono osservati e controllati durante lo svolgimento del proprio impiego (ivi, 365). La prima tra le due varianti sopracitate, tuttavia, prevede un'estrema fiducia reciproca tra cliente e post-editor, sebbene non di rado il cliente

fornisca precedentemente una stima delle ore che dovrebbero essere necessarie e sufficienti per effettuare il post-editing; ciononostante, è anche la forma più equa tra le due, in quanto permette di tener conto delle debolezze presentate dall'output, la cui qualità in certi casi non è ottima, quindi anche delle difficoltà incontrate dal post-editor nello svolgimento del lavoro (ibidem).

Sara Grizzo fornisce, infine, un'ulteriore riflessione sul post-editing della traduzione automatica e sulla revisione di un *fuzzy match*: l'autrice si chiede, infatti, se abbia ancora senso differenziare i due processi (di post-editing e di revisione), visto il netto miglioramento della qualità dell'output della MT, dovuto all'utilizzo dei sistemi di traduzione neurale, e se gli strumenti CAT siano da considerare conformi ai tempi attuali oppure no (ibidem). Aggiunge che il pagamento su base oraria risulterebbe sicuramente il più sensato, e dunque il migliore, se, a un certo punto, le attuali soluzioni software, con le correnti tariffe annesse, venissero effettivamente abbandonate e la traduzione fosse effettuata liberamente, servendosi del supporto di strumenti diversi (ibidem).

### **3.2.3 Pagamento *edit-distance***

Un terzo metodo di remunerazione, che si differenzia ulteriormente dai precedenti due, è il pagamento secondo la distanza di modifica (*edit-distance*). Il concetto di “distanza di modifica” risale agli anni '50 del secolo scorso e venne introdotto dallo scienziato russo Levenshtein (per tale ragione la distanza di modifica è anche conosciuta con il nome di “distanza di Levenshtein”) come metodo di calcolo del minimo sforzo richiesto per la trasformazione di una stringa in un'altra, in base alle somiglianze e alle differenze tra due specifiche stringhe (Asmuth, 2020:370). Esistono diverse metriche per il calcolo della distanza di modifica, ma le due più conosciute sono proprio la distanza di Levenshtein e TER(p), sebbene tutte funzionino nello stesso modo (Schelfhout, 2022). Le operazioni richieste, per la trasformazione di una frase in un'altra, possono essere di diverso tipo, quali operazioni di inserimento, eliminazione o sostituzione (Grizzo, 2020:365). Ad esempio, per trasformare la parola inglese *cat* nella parola *hats* sono richieste due operazioni, ovvero è necessario cambiare la lettera “c” rendendola “h” e infine aggiungere la lettera “s” alla fine della parola (Asmuth, 2020:370). All'interno dell'industria della traduzione esistono, oramai da decenni, diversi tipi di metriche di calcolo della distanza di modifica, come ad esempio BLEU, usati per la valutazione

della qualità dell'output dei sistemi di traduzione automatica, attraverso il confronto tra l'output stesso e una traduzione di riferimento, prodotta da un traduttore umano (ibidem). In genere, la quantità di modifiche apportate non rappresenta un numero totale: il numero di modifiche è spesso diviso per la lunghezza della stringa (ad esempio, una stringa può essere lunga tre parole come tredici) e, in base a tale suddivisione, si crea un punteggio finale del calcolo della distanza di modifica, di solito compreso tra 0 e 1, il quale rappresenta la percentuale di caratteri (o parole, a seconda della metrica di misura utilizzata), modificati in una frase (ivi, 371; Schelfhout, 2022). Questa operazione si chiama "normalizzazione" e rende i valori della distanza di modifica più comparabili tra stringhe di lunghezza diversa (Schelfhout, 2022).

Il calcolo della distanza di modifica compara l'output di un sistema MT con una traduzione di riferimento senza contesto (Asmuth, 2020:371): anche se semanticamente la frase non è corretta, la traduzione automatica neurale raggiungerà sempre un buon punteggio perché risulta linguisticamente fluente nell'output che crea (ibidem). Per tale ragione, nell'effettuare l'analisi della qualità di un sistema MT, devono essere utilizzati anche feedback, misurazioni di efficienza e altri sistemi di valutazione basati sull'uomo, accanto al calcolo della distanza di modifica (ibidem).

La distanza di modifica può anche essere utilizzata per valutare lo sforzo compiuto dal post-editor: infatti, a livello teorico, un numero minore di modifiche apportate equivale a uno sforzo minore compiuto, e dunque, il post-editor dovrebbe essere pagato meno, solamente per il lavoro effettivamente svolto (Schelfhout, 2022). Tuttavia, questo ragionamento non è totalmente corretto, in quanto il lavoro del post-editor è composto da una lunga serie di fasi: l'assunzione del lavoro; la lettura di un segmento nella lingua sorgente; la lettura del segmento di arrivo e la sua comparazione con quello sorgente; la lettura dei precedenti e seguenti segmenti per la comprensione del contesto; il controllo della lista dei termini (in alcuni casi, in cui vi è una necessità di esattezza dei termini usati, è necessario controllare che tutti siano stati inseriti nel database terminologico) e dello stile; il post-editor deve effettuare ulteriori ricerche riguardo al testo e al suo contenuto, pensando anche al miglior approccio da attuare; finalmente arriva la parte di traduzione o modifica, in cui avviene, per la prima volta, la scrittura vera e propria; il post-editor deve ricontrollare il lavoro e la sua qualità, e, infine, consegnare il lavoro svolto al cliente (ibidem). Se l'output della traduzione automatica è già buono in

partenza, ovviamente è necessario un numero di battute relativamente basso (ibidem). Vi è una concezione errata alla base, ovvero quella di valutare il post-editing (ma anche la traduzione in generale) come un mero lavoro di digitazione: lo sforzo richiesto dal post-editor è, tuttavia, una combinazione dell'esperienza lavorativa, della qualità dell'output e delle aspettative del cliente sul prodotto finale (ibidem). La distanza di modifica, dunque, può misurare solo i caratteri modificati e non può tener conto del calcolo dello sforzo totale impiegato (ibidem).

Asmuth (2020:372), riporta un esempio di come venga calcolato il prezzo del post-editing sulla base del calcolo della distanza di modifica presso SAP, un'azienda multinazionale tedesca, che usa CharacTER come metrica di calcolo della distanza di modifica. Per la trasformazione del punteggio ottenuto dal calcolo della distanza di modifica in un sistema di pagamento, l'azienda applica uno sconto sul prezzo per parola per la nuova traduzione senza un match nella TM (ibidem). Lo sconto minimo equivale allo 0%, ovvero pagata come una traduzione nuova, senza alcuna differenza, quando la frase deve essere completamente scritta e non vi è alcun match presente; non esiste uno sconto del 100% nel caso, in cui la traduzione non abbia bisogno dunque di modifiche, perché viene compreso che l'output della MT deve comunque essere letto e controllato più volte rispetto a un match TM (ibidem). Viene riportata una tabella esemplificativa:

<b>EDC Score</b>	<b>Discount on price for new translation</b>
0-0.19	50%
0.2-0.39	40%
0.4-0.59	30%
0.6-0.79	20%
0.8-0.99	10%
1	0%

Tabella 3: calcolo del prezzo in base alla distanza di modifica per SAP (Asmuth, 2020:372)

La compensazione in base alla distanza di modifica, secondo Asmuth, è sicuramente una metodologia migliore rispetto a quella per parola, sebbene difficilmente quest'ultimo metodo verrà sostituito nel corso degli anni (ivi, 373). L'autore aggiunge che, secondo la propria opinione, una buona soluzione per una compensazione equa potrebbe essere l'uso di una retribuzione su base oraria supportata da meccanismi di

controllo come la distanza di modifica, fornendo, allo stesso tempo, un incentivo alla produzione di un output della miglior qualità possibile (ibidem).

### **3.2.4 Direttive della società TAUS sulla compensazione**

Il TAUS non fornisce indicazioni sui diversi metodi di pagamento, ma indica delle linee guida rivolte ai post-editor, per la determinazione di un metodo di pagamento adatto alle proprie esigenze. Infatti, viene ribadito che non esiste un sistema unico per la determinazione del pagamento del post-editing, ma, per l'appunto, è necessario fare riferimento a una combinazione di approcci (TAUS, 2017:33). La società TAUS, inoltre, nomina come unici due metodi di pagamento per il post-editing quello a parola e quello su base oraria (ibidem).

Qualsiasi combinazione di approcci si scelga, viene evidenziata la necessità di creare un modello che sia predittivo, equo e appropriato:

- predittivo: il modello scelto deve aiutare a prevedere il costo del post-editing, stabilendo il prezzo in anticipo (ivi, 34). Il modello, dunque, dovrebbe essere in grado di calcolare il costo di un determinato volume di testo in maniera istantanea (ibidem);
- equo: un modello per la determinazione del prezzo del MTPE dovrebbe essere prima di tutto equo e tutte le parti coinvolte nel processo di traduzione (ovvero clienti, fornitori di servizi linguistici, traduttori, ecc.) dovrebbero parteciparvi, concordando un modello che rifletta lo sforzo richiesto e permettendo una misurazione affidabile del lavoro (ibidem). Inoltre, ai traduttori deve essere fornita una formazione sul post-editing e devono essere fissate aspettative realistiche (ibidem);
- appropriato: il modello scelto deve tenere in considerazione le caratteristiche del contenuto (ivi, 35). Infatti, in base a quest'ultimo, la qualità dell'output della MT varia notevolmente, e i livelli di post-editing richiesti possono differire molto (ibidem). Anche la coppia linguistica coinvolta nel processo di traduzione influisce sulla qualità dell'output e, di conseguenza, i prezzi possono variare in base alle due casistiche precedentemente elencate (ibidem). La qualità della MT dovrebbe essere stabilita in un'unica valutazione globale, prima dell'adozione della traduzione automatica, effettuata per ogni sistema e tipo di contenuto (ibidem).

Risulta importante ricordare che sia il metodo di valutazione della qualità sia la conseguente determinazione dei prezzi devono essere trasparenti (ibidem). Al fine di ottenere un modello predittivo, equo e appropriato, è necessario combinare diversi approcci. È dunque raccomandato combinare tre approcci in particolare:

- un metodo di valutazione automatico (ad esempio, GTM, TER, BLEU oppure l'analisi inversa della traduzione automatica – *MT Reversed Analysis*) (ivi, 36). La metrica GTM (*General Text Matching*) misura le somiglianze tra l'output prodotto dal sistema di traduzione automatico e le traduzioni di riferimento prodotte dall'uomo, calcolando la distanza di modifica (ivi, 37). Invece, con “analisi inversa della MT” si intende un approccio che mira a correlare la qualità dell'output di una MT con i *fuzzy match*: calcola il livello di *fuzzy-match* dei segmenti prodotti dalla traduzione automatica, rispetto ai segmenti di cui si è effettuato il post-editing (ibidem). Per questo approccio si usa il modello della determinazione del prezzo già esistente per gli strumenti CAT (ibidem).
- una revisione umana della qualità dell'output: essa può essere usata per valutare la qualità dell'output prodotto dalla traduzione automatica, per controllare se la qualità finale del contenuto, dopo il post-editing, è del livello desiderato (ivi, 38);
- una valutazione della produttività: consiste nel calcolo della differenza di velocità tra il tempo necessario per la traduzione effettuata da zero e il post-editing dell'output della MT (ibidem). I risultati possono cambiare di volta in volta, a seconda del sistema utilizzato (ibidem).

Combinando questi approcci ogni volta che si utilizza un nuovo sistema di traduzione automatica, il post-editor possiede la certezza di avere a propria disposizione un modello di pagamento predittivo, equo e appropriato (ivi, 39).

### **3.3 Metodologia: associazione libera**

Come precedentemente sottolineato, la prima parte del questionario è stata progettata usando la tecnica dell'associazione libera. Quella dell'associazione libera è una tecnica usata soprattutto per raccogliere delle percezioni sociali e rappresentarle (Dany *et al.*, 2014:489). Le rappresentazioni sociali sono definite come “a specific form of knowledge – common knowledge – whose contents show the operation of generative

processes and socially marked functions<sup>15</sup>” (Jodelet, 1984:36, in Brondi & Neresini, 2018:2). Un punto chiave delle rappresentazioni sociali è il ruolo fondamentale dei processi generativi, denominati “ancoraggio” e “oggettivazione” da Moscovici (1984), e attraverso i quali alcuni concetti, normalmente non familiari, lo diventano (Brondi & Neresini, 2018:2). L’ancoraggio consiste nell’applicazione a un nuovo fenomeno sociale di un concetto preesistente, mentre l’oggettivazione è la trasformazione di un fenomeno sociale in un oggetto concreto e tangibile (Wagner & Hayes, 2005, in Brondi & Neresini, 2018:2). Le rappresentazioni sociali richiedono continui processi di familiarizzazione, in quanto soggetti a diverse sfide sociali (Brondi & Neresini, 2018:2). Le rappresentazioni sociali possono essere individuate tramite l’utilizzo della tecnica dell’associazione libera.

Nell’associazione libera, la procedura da seguire deve rispettare un preciso ordine: innanzitutto vengono fornite al partecipante una o più parole-stimolo, ovvero delle parole che rappresentano l’oggetto centrale dell’indagine, a partire dalle quali devono produrre altre parole o espressioni che spontaneamente vengono loro in mente (ivi, 490). Questo metodo permette di accedere alle cognizioni per descrivere le rappresentazioni sociali, e hanno lo scopo di rivelare tracce della memoria collettiva e di ragionare sulla struttura di quest’ultime (Lahlou & Abric, 2011; Flament & Rouquette 2003:58, in Dany *et al.*, 2014:490). L’associazione libera presenta alcuni vantaggi, in quanto permette di processare direttamente i dati derivanti dalle espressioni libere dei singoli individui, è veloce da applicare e analizzare e può essere utilizzata facilmente dai partecipanti, che ne comprendono il suo uso senza problemi (Abric 2003; Moliner *et al.* 2002, in Dany *et al.*, 2014:490).

Dany *et al.* (2014:490) elencano quattro principali tecniche che si basano sull’incrocio (tabulazione incrociata) di due criteri (Jodelet 1972; Flament e Rouquette 2003): il primo è un criterio numerico, basato sul fatto che un soggetto è invitato a produrre un numero  $x$  di parole, che normalmente, in questo tipo di studi, sono tre o cinque; il secondo è un criterio di limitazione, ovvero possono essere presenti delle restrizioni di tipo grammaticale (ad esempio, è necessario rispondere utilizzando solo aggettivi, sostantivi, verbi, ecc.), semantico (ad esempio, rispondere con un termine che

---

<sup>15</sup> “Una forma specifica di conoscenza – di conoscenza comune - i cui contenuti mostrano il funzionamento dei processi generativi e delle funzioni socialmente marcate” (tradotto da me).

descrive un sentimento; Campos e Rouquette 2000 in Dany *et al.*, 2014:490), o di natura sostitutiva (ad esempio, rispondendo come farebbe una diversa categoria di persone, ovvero immedesimandosi; Guimelli e Deschamps 2000; Marie et al. 2010; Lo Monaco et al. 2009 in Dany *et al.*, 2014:490). Dall'incrocio dei due criteri sopracitati nascono, dunque, quattro diverse tecniche fondamentali di associazione libera (Flament e Rouquette 2003 in Dany *et al.*, 2014:490): l'associazione libera a risposta singola (solo una parola, indotta senza restrizione), l'associazione controllata a risposta singola (solo una parola indotta con restrizione), l'associazione libera a risposta multipla (più parole indotte senza restrizione) e l'associazione controllata a risposta multipla (più parole indotte con restrizione). Tra queste quattro tipologie, l'associazione libera a risposta multipla è la più apprezzata (Dany *et al.*, 2014:490).

L'associazione libera aiuta a fornire una rappresentazione sociale. In particolare, questo approccio ipotizza che la rappresentazione sociale sia un sistema organizzato gerarchizzato, composto a sua volta da due sottosistemi, ovvero un sistema centrale e uno periferico, che interagiscono tra di loro (ibidem). Un numero limitato di elementi fa parte del sistema centrale, i quali costituiscono la base consensuale e comune della memoria collettiva, oltre al sistema di norme a cui un gruppo fa riferimento (Abric, 2001, in Dany *et al.*, 2014:490). Questo sistema svolge una serie di funzioni, quali la generazione del significato della rappresentazione, e determina la sua organizzazione, oltre a incoraggiare il consenso e l'omogeneità culturale e psicologica di un gruppo, o di un singolo individuo: il sistema centrale è il "vero fondamento della struttura e natura della rappresentazione" (Abric, 1987:68, in Dany *et al.*, 2014:491). Lo status centrale di un elemento si estende anche alla sua dimensione qualitativa, non solo quantitativa: infatti, la sua centralità è definita dal significato che l'elemento dà alla rappresentazione (Moliner, 1994, in Dany *et al.*, 2014:491). Alcuni termini possono, naturalmente, presentarsi con una certa frequenza dal punto di vista quantitativo all'interno delle risposte dei partecipanti (Moliner, 1994, in Dany *et al.*, 2014:491): si tratta della cosiddetta "comunità di opinione" (*community of opinion*). Normalmente si potrebbe pensare che un gruppo ideale sia quello che condivide le stesse idee (Dany *et al.*, 2014:491). Tuttavia, anche un gruppo di partecipanti, le cui risposte si discostano da quelle della maggioranza, può essere considerato come un gruppo ideale, pur formando un sistema periferico (ibidem). Fanno parte del sistema periferico tutti quegli elementi

che forniscono un'espressione concreta e illustrano la rappresentazione, sebbene siano più diversificati e flessibili degli elementi del sistema centrale (ibidem).

Una volta raccolto il corpus ottenuto durante il compito di associazione libera, i dati devono essere processati. Esistono due metodi per effettuare questo processo: *rank-frequency method* (grado di frequenza) e *ranked association method* (associazioni classificate) (ibidem). Il primo dei due, basato sul “grado di frequenza”, è il meno recente, e considera due indicatori, ovvero la frequenza di un elemento o di parole indotte (dunque, la dimensione collettiva, quantitativa) e la sua apparizione spontanea nella classifica (identificato come un elemento di un'operazione individuale in quanto il partecipante usa una determinata parola in base alla disponibilità cognitiva) (ibidem). Vi è una diretta relazione tra i due criteri, e la loro congruenza può costituire un indicatore della centralità di un elemento (Abric, 1994, in Dany *et al.*, 2014:491).

La tabulazione incrociata di tali criteri permette di creare una tabella a quattro celle che rappresenta le quattro zone distinte della rappresentazione (Tabella 4): nella prima cella si trovano i termini che sono citati per primi e più frequentemente, dunque i termini più salienti e significanti per i partecipanti, che hanno la più alta probabilità di appartenere al nucleo centrale della rappresentazione; nella quarta cella sono contenuti i termini meno frequentemente citati e quelli citati per ultimi, che rappresentano il sistema periferico; le celle 2 e 3, invece, sono delle zone di potenziale cambiamento, in cui vengono date delle informazioni contraddittorie (ad esempio, le parole potrebbero essere state citate per prime ma la loro frequenza è bassa), fornendo lo spunto per dei

Frequency	Appearance ranking	
	Low	High
High	Cell 1—core zone	Cell 2—potential change zone
Low	Cell 3—potential change zone	Cell 4—periphery

Tabella 4: Metodo *rank-frequency* (Dany *et al.*, 2014:491)

Tuttavia, questa metodologia è stata spesso messa in discussione, in quanto si presuppone che i primi termini scritti siano anche i più importanti: secondo alcune ricerche ciò può essere considerato veritiero (Jones, 1958; Winnick & Kressel, 1965), ma secondo altri studiosi, come Abric (2003:53, in Dany *et al.*, 2014:492), le informazioni “essenziali” sarebbero quelle venute in mente al partecipante solo dopo un

momento di “riscaldamento”, e non quelle pensate e scritte di getto. Dunque, per venire incontro alle discussioni sull’argomento, Abric (2003), propose di sostituire la classificazione per apparizione con la classificazione per importanza, chiedendo ai partecipanti di elencare i termini in ordine di importanza (in Dany *et al.*, 2014:492). Questo nuovo metodo prende il nome di *ranked associations* (associazioni classificate), e la tabulazione incrociata dei due nuovi elementi (ovvero la frequenza e la graduatoria di importanza), danno luogo a quattro celle (Tabella 5): la prima uguale alla precedente, ovvero contenente gli elementi chiave, centrale; la seconda cella (prima periferia) contiene termini importanti ma poco frequenti; la cella 3 rappresenta gli elementi contrastanti, ovvero quegli elementi nominati da parte di poche persone ma considerati da queste come molto importanti; infine la cella 4 rappresenta la seconda periferia, contenente elementi poco frequenti e poco importanti (ivi, 492-493). Pur mostrando delle somiglianze, i due metodi sono basati su dei presupposti epistemologici distinti, che conferiscono uno status diverso agli elementi della rappresentazione prodotta (ivi, 493).

Frequency	Importance	
	High	Low
High	Cell 1—core zone	Cell 2—first periphery
Low	Cell 3—contrasting elements	Cell 4—second periphery

Tabella 5: Metodo *ranked associations* (Dany *et al.*, 2014:492)

### 3.4 Prima sezione: percezione del post-editing e della sua retribuzione

Nel corso delle seguenti sezioni verranno analizzati i risultati ottenuti dal questionario sottoposto ai 132 partecipanti. Come anticipato, il questionario si presenta suddiviso in due parti principali: la prima, dedicata al compito di associazione libera, rivolta a tutti i partecipanti; la seconda, diretta ai post-editor, al fine di comprendere la loro opinione sul tema della compensazione nel post-editing. Per quanto riguarda il compito di associazione libera, esso è basato sul grado di frequenza: infatti, non è stato chiesto ai partecipanti di menzionare i termini in ordine di importanza. Successivamente sono state proposte sia una serie di domande aperte che chiuse, a risposta singola o multipla, in entrambe le parti costituenti il questionario.

Innanzitutto, è stata svolta l’analisi dalla percezione del post-editing, sia da parte dei post-editor che del pubblico generale. Il presupposto di partenza, antecedente

all'elaborazione dei dati, è che il post-editing non sia un lavoro particolarmente apprezzato tra i traduttori ed esperti del settore, ma che, nonostante tutto, debba essere effettuato perché vi è un'enorme richiesta di post-editing all'interno del mercato della traduzione. Inoltre, l'aspettativa è che vi sia una scarsa conoscenza del tema all'interno del pubblico generale.

### **3.4.1 Compito di associazione libera**

Nel compito di associazione libera sono state poste all'attenzione dei partecipanti due parole-stimolo: "traduzione" e "post-editing". A partire dalle due parole-stimolo, è stato richiesto ai partecipanti di produrre dalle due alle cinque parole, che fossero aggettivi o sostantivi, associate a loro. I risultati ottenuti sono stati numerosi e vari, registrando, per ognuna delle parole-stimolo, più di 200 termini differenti. Tuttavia, sono state mantenute solamente le risposte ritenute adatte (dunque sono stati scartati verbi, o altre formulazioni che non rispettassero la consegna). In modo particolare, sono state considerate le parole con un numero maggiore di occorrenze, ovvero quelle citate da più partecipanti, e le parole che, sebbene siano state scritte anche da un solo partecipante, sono state ritenute più interessanti ai fini della ricerca. Le parole stimolo sono state proposte in lingua inglese, come il resto del questionario, e per questo la maggioranza dei risultati ottenuti sono in lingua inglese: vengono dunque riportati in lingua inglese con la corrispondente traduzione italiana accanto. Il numero di occorrenze è indicato tra parentesi, subito dopo il sostantivo o aggettivo a cui fa riferimento.

La prima parola-stimolo sottoposta all'attenzione dei partecipanti è stata *translation*, "traduzione". La frequenza è rappresentata all'interno della Tabella 6: le parole all'interno della prima cella rappresentano il nucleo centrale, contenente i termini maggiormente nominati, ovvero *language(s)* (42), "lingua/e", sia nella versione singolare che plurale del sostantivo, e *communication* (22), "comunicazione". Si tratta, in realtà, di parole molto generiche, e per tale ragione nominate da molti partecipanti, in modo particolare da coloro che non hanno alcuna esperienza nel campo della traduzione. Successivamente, nella seconda cella, ovvero la prima zona di potenziale cambiamento, vi sono termini quali i sostantivi *culture(s)* (14), "cultura/e", *text(s)* (11), "testo/i", entrambi sia nella loro forma singolare che quella plurale anche in questo caso, e *understanding* (12), "comprensione". Alla terza cella appartengono parole come *interpretation* (10), "interpretazione", *meaning* (8), "significato", *accuracy* (6) e

*accurate* (3), “accuratezza” e “accurato”. Infine, nella quarta cella, ovvero la “periferia”, rientrano i termini meno citati, nominati solamente un una o due volte: *fun* (2), “divertente”, e *concentration* (1), “concentrazione”, ne sono un esempio.

È possibile raggruppare le parole citate dai partecipanti anche in base alle somiglianze tra loro esistenti. Infatti, vi sono dei termini che rappresentano le caratteristiche della traduzione, come alcune parole legate all’ambito della creatività come *creativity* (5), *creation* (3) e *creative* (3), “creatività”, “creazione” e “creativo”, aggettivi quali *difficult* (4), “difficile”, *cultural* (4), “culturale”, oppure ancora *knowledge* (4), “conoscenza”, *adaptation* (5), “adattamento”, *training* (1), “allenamento”, *professional* (3), “professionale”, *skill(s)* (4), “abilità”. Quelli precedentemente nominati, sono tutti termini che forniscono un’indicazione sulla percezione, soprattutto da parte di traduttori/post-editor (o eventualmente da parte di studenti di materie linguistiche), del lavoro di traduzione, che viene normalmente inteso come un lavoro che richiede particolare precisione, presentando, dunque, anche delle difficoltà, degli ostacoli da affrontare. Quello di traduzione può essere definito come un lavoro creativo, che richiede determinate conoscenze e capacità come prerequisiti, e, di conseguenza, non può essere svolto da chiunque, bensì da esperti nel settore, in seguito a un’adeguata formazione nell’ambito. La traduzione è, poi, strettamente legata alla cultura, poiché prevede la conoscenza culturale oltre che linguistica di ciò che si deve tradurre, affinché i riferimenti culturali presenti vengano chiariti. Alcuni traduttori, partecipanti allo studio, hanno definito la traduzione come un ponte tra culture diverse, espresso dal sostantivo *bridge* (6), “ponte”.

Altri sostantivi e aggettivi citati sono legati al lavoro di traduzione e al suo svolgimento: *written* (6), “scritto”, e *writing* (4), “scrittura”, *transfer* (6), “trasferimento”, inteso come il trasferimento di significato da una lingua a un’altra, *foreign language* (6), “lingua straniera”, *word(s)* (5), “parola/e”, *mediation* (4), “mediazione”, *message* (4), “messaggio”, *context* (3), “contesto”, *dictionary* (3) e *vocabulary* (3), “dizionario” e “vocabolario”, *literal* (2), “letterale”. Infatti, il lavoro di traduzione viene spesso associato alla scrittura. Il traduttore, inoltre, è un mediatore tra lingue straniere e culture, e il suo compito è anche quello di riuscire a trasferire un determinato messaggio da una lingua verso un’altra. I traduttori e post-editor hanno associato alla parola “traduzione” i sostantivi *work* (4) o *job* (3), o ancora *profession* (2),

“lavoro”, “professione”, in alcuni casi accompagnati dall’aggettivo possessivo *my* (3), “mio”, con l’intento di sottolineare che si tratta del proprio lavoro. Un traduttore, infatti, ha indicato anche la parola *passion* (1), “passione”, evidenziando ulteriormente la sua vicinanza a tale professione.

Vi sono una serie di aggettivi, citati da alcuni partecipanti, che hanno una connotazione particolarmente positiva: il lavoro di traduzione viene definito *rewarding* (3), “appagante”, *interesting* (3), “interessante”, *original* (3), “originale”, *useful* (3), “utile”, *fun* (2), “divertente”, *precise* (2), “precisa”, *specialised* (2), “specializzata”. *faithful* (1), “fedele”, *vital* (1), “vitale”, *fast* (1), “veloce”, *fulfilling* (1), “appagante”, *essential* (1), “essenziale”, *good* (1), “buono”, *reliable* (1), “affidabile”, *best fit* (1) “il più adatto”. Inoltre, sono stati citati alcuni sostantivi, anch’essi contenenti una connotazione positiva: *precision* (2), “precisione”, *ability* (1), “abilità”, *value* (1), “valore”, *trust* (1), “fiducia”, *sense* (1), “senso”, *respect* (1), “rispetto”, *proficiency* (1), “competenza”, *professionalism* (1), “professionalità”, *freedom* (1), “libertà”, *excellence* (1), “eccellenza”, *faithfulness* (1), “fedeltà”, *specialisation* (1), “specializzazione”, *life* (1), “vita”. Gli aggettivi e sostantivi in questione descrivono il lavoro di traduzione in modo molto positivo, in quanto permette di ottenere un risultato finale eccellente e di senso compiuto. Sebbene alcune parole siano state nominate solamente da un partecipante o due, si rivelano particolarmente utili per fornire un’idea più completa della percezione della traduzione sia tra il pubblico generale che tra i post-editor stessi.

Ciononostante, la traduzione nasconde delle insidie e, per tale motivo, alcuni partecipanti hanno nominato aggettivi e sostantivi relativi alla difficoltà di questa professione: infatti, la traduzione viene vista, in alcuni casi, come una sfida e definita, dunque, *challenging* (3), “impegnativa” e *arduous* (1), “faticosa”, in quanto richiede “concentrazione”, *concentration* (1). Un partecipante la descrive come *time-consuming* (1), poiché, secondo la sua opinione, richiede molto tempo. Viene definita anche *lost* (1), “perduta”, e *ignored* (1), “ignorata”: infatti, oggi giorno all’interno del mercato della traduzione vi è una sempre maggiore richiesta di post-editing della traduzione automatica. Tale cambiamento è motivo di preoccupazione per i traduttori, che temono di essere surclassati dalle macchine. La componente umana, all’interno delle risposte ottenute, è comunque presente: è stato nominato l’aggettivo *human* (2), “umano”, e il

sostantivo *people* (3), “persone”. Tuttavia, vi è anche un riferimento alla traduzione automatica: *machine translation* (3).

Infine, sono presenti solamente due parole relative alla sfera del pagamento, ovvero *money* (1), “soldi”, e *decent pay* (1), “retribuzione soddisfacente”. Sebbene siano solamente due termini e dunque non possano essere considerati validi per l’intero gruppo di partecipanti, sono indicativi di una situazione positiva, apprezzabile, in quanto sembra trasparire che il lavoro di traduzione permetta di ottenere dei buoni guadagni.

Frequenza	Grado	
	Alto	Basso
Alto	language(s) (42) communication (22)	culture(s) (14) understanding (12) text(s) (11)
Basso	interpretation (10) meaning (8) accuracy (6) bridge (6) transfer (6) written(6) foreign language(s) (6) adaptation (5) creativity (5) word(s) (5) cultural (4) difficult (4) knowledge (4) mediation (4) message (4) skill(s) (4) work (4) writing (4)	faithfulness (2) fun (2) Google (2) human (2) literal (2) precise (2) precision (2) profession (2) specialized (2) ability (1) arduous (1) best fit (1) concentration (1) decent pay (1) essential (1) excellence (1) faithful (1) fast (1)

accurate (3)	freedom (1)
challenging (3)	fulfilling (1)
context (3)	good (1)
creation (3)	ignored (1)
creative (3)	life (1)
dictionary (3)	lost (1)
interesting (3)	money (1)
job (3)	passion (1)
machine translation (3)	professionalism (1)
my (job/work) (3)	proficiency (1)
original (3)	reliable (1)
people (3)	respect (1)
professional (3)	sense (1)
rewarding (3)	specialization (1)
useful (3)	time-consuming (1)
vocabulary (3)	training (1)
	trust (1)
	value (1)
	vital (1)

Tabella 6: Rappresentazione di "translation"

La seconda parola-stimolo sottoposta all'attenzione dei partecipanti allo studio è stata la parola "post-editing". La frequenza è rappresentata all'interno della Tabella 7: all'interno della prima cella, rappresentante il nucleo centrale, vi è il termine *correction(s)* (21), "correzione/i". Nella seconda cella, corrispondente alla prima zona di potenziale cambiamento, rientra la parola *revision(s)* (14), "revisione/i". Vi sono, poi, una serie di sostantivi e aggettivi, particolarmente nominate dai partecipanti, facenti parte della terza cella come, ad esempio, *review* (9), "revisione", *translation* (7), "traduzione", e *automatic* (4), "automatico". Infine, alla zona periferica appartengono i termini che presentano una sola occorrenza, quali *chaos* (1) e *hate* (1), "odio".

Anche in questo caso è possibile raggruppare le parole citate dai partecipanti in base alle somiglianze tra loro esistenti: vi sono delle parole strettamente legate al lavoro di post-editing come *check* (9), "controllo" *machine translation* (8), "traduzione

automatica”, *editing* (6), “correzione”, *adjustment(s)* (5), “adattamento”, *error(s)* (5), “errore/i”, *AI* (4), “intelligenza artificiale”, *challenging* (3), “impegnativo”, *language(s)* (3), “lingua/e”, *machine* (3), “macchina”, *modification(s)* (3), *computer* (2), *mechanical* (2), “meccanico”, *attention* (2), “attenzione”, *skill(s)* (2), “abilità”, *reading* (2), “lettura”. Il post-editing, infatti, si basa sulla revisione e modifica dell’output prodotto dalla traduzione automatica, che è, a sua volta, una tipologia di intelligenza artificiale (AI). Sebbene possa sembrare un procedimento semplice, contiene una serie di complessità, e per tale ragione è considerato un lavoro impegnativo: infatti, spesso risulta particolarmente difficile individuare tutti gli errori prodotti da un sistema neurale. L’aggettivo *human* (2), “umano”, è un riferimento alla componente umana presente e necessaria nel post-editing. Vi sono, poi, dei riferimenti ai due sistemi di traduzione automatica online e gratuiti, oggi maggiormente usati, quali Google Translate (3) e DeepL (2).

Alcuni partecipanti hanno descritto il post-editing utilizzando una serie di aggettivi e sostantivi positivi. Ne sono un esempio: *modern* (4), “moderno”, “economico”, *fast* (3), “veloce”, *speed* (3), “velocità”, *precision* (3), “precisione”, *improvement* (3), “miglioramento”, *quick* (3), “rapido/breve”, *easy* (2), “facile”, *interesting* (1), “interessante”, *helpful* (1), “utile”, *indispensable* (1), “indispensabile”, *innovative* (1), “innovativo”, *precise* (1), “preciso”, *time-saving* (1), “che fa risparmiare tempo”, *well-done* (1), “ben fatto”, espressioni riguardo all’efficienza come *efficiency* (1) e *efficient* (1). Infatti, il lavoro di post-editing è ritenuto utile da diversi partecipanti, i quali ripongono estrema fiducia nel futuro dell’intelligenza artificiale; inoltre, secondo molti è un procedimento veloce, che permette di risparmiare molto tempo, pur mantenendo la precisione.

Sono altrettanto presenti, tuttavia, numerosi aggettivi e sostantivi con una connotazione altamente negativa nei confronti del post-editing: infatti, il post-editing viene definito particolarmente *boring* (3), o *tedious* (2), ovvero “noioso”, *difficult* (2), “difficile”, *exhausting* (2), “sfiancante”, *uncreative* (2), “non creativo”, *time consuming* (2), “dispendioso” (in termini di tempo), *tricky* (2), “complicato”, *restricted* (2), “limitato”, *stupid* (1), “stupido”, *superficial* (1), “superficiale”, *frustrating* (1), “frustrante”, *uninspiring* (1), “non stimolante”, *worrisome* (1), “preoccupante”, *dreary* (1), “monotono”, *unsuitable* (1), “inadeguato”, *not worth doing* (1), “non ne vale la

pena”. Inoltre, parole come *low quality* (3) e *poor quality* (1) indicano una qualità dell’output della MT particolarmente scarsa. Alcuni partecipanti utilizzano altre parole per indicare i punti negativi del post-editing, quali *mistakes* (2), “errori”, *chaos* (1), *exploitation* (1), “sfruttamento”, *hate* (1), “odio”, *hurry* (1), “fretta”, *hindrance* (1) “ostacolo”, *inaccuracy* (1), “inesattezza”, *monotony* (1), “monotonia”, *nightmare* (1), “incubo”, *pain* (1), “dolore”, *slave* (1), “schiavo”, *not real translation* (1), “non una traduzione reale”. Si nota, dunque, un’avversità verso il post-editing: non è inteso come una vera traduzione, ma come un’attività particolarmente noiosa, oltre che complessa, in quanto nasconde delle insidie. Viene giudicato come un lavoro frustrante e limitante, dispendioso in termini di tempo e poco stimolante. Gli elementi negativi superano quelli positivi, dimostrando che il post-editing non è un’attività apprezzata da molti traduttori.

Alcuni dei termini elencati dai partecipanti, anche in questo caso, sono legati precisamente alla sfera del pagamento e dei costi da sostenere, che è anche oggetto dello studio in questione. Alcuni sembrano avere una connotazione positiva, quali *cheap* (3), “economico”, o altri che indicano un risparmio sui come *cost-saving* (1) e *money saving* (1) “risparmio dei costi”, o ancora *cost-effectiveness* (1), “efficacia dei costi” e *low-cost* (1), “a basso costo”. Altri, tuttavia, sembrano avere una connotazione molto negativa, quali *underpaid* (2), “sottopagato” e *underpayment* (1), “sottopagamento”. Una serie di parole è preceduta dall’aggettivo *low*, “basso”: *low fare* (1), “tariffa bassa”, *low payment* (1), “pagamento basso”, *lower pay* (1), “retribuzione più bassa”.

Frequenza	Grado	
	Alto	Basso
Alto	correction(s) (21)	revision(s) (14)
Basso	check (9) review (9) machine translation (8) translation (7) editing (6) adjustment(s) (5) error(s) (5) AI (4) automatic (4)	chaos (1) cost-effectiveness (1) cost-saving (1) dreary (1) efficiency (1) efficient (1) exploitation (1) frustrating (1)

modern (4)	hate (1)
boring (3)	helpful(1)
challenging (3)	hindrance (1)
cheap (3)	hurry (1)
fast (3)	inaccuracy (1)
Google translate (3)	indispensable (1)
Improvement (3)	innovative (1)
language(s) (3)	interesting (1)
low quality (3)	low fare (1)
machine (3)	low payment (1)
modification(s) (3)	low-cost (1)
precision (3)	lower pay (1)
quick (3)	money saving (1)
speed (3)	monotony (1)
attention (2)	nightmare (1)
computer (2)	not real translation (1)
DeepL (2)	not worth doing (1)
Difficult (2)	pain (1)
easy (2)	poor quality (1)
exhausting (2)	precise (1)
human (2)	rough (1)
mechanical (2)	slave (1)
mistakes (2)	stupid (1)
reading (2)	superficial (1)
restricted (2)	time saving (1)
skill(s) (2)	uncreative (1)
tedious (2)	under-payment (1)
time-consuming (2)	uninspiring (1)
tricky (2)	unsuitable (1)
underpaid (2)	well-done (1)
	worrisome (1)

Tabella 7: Rappresentazione di "post-editing"

È evidente una netta differenza tra i termini associati alla traduzione e quelli invece legati al post-editing. Infatti, la maggioranza delle parole riguardanti l'ambito della traduzione hanno delle sfumature positive: sebbene sia particolarmente difficile e complesso, in quanto è necessario avere un'ottima conoscenza linguistica e culturale, quello del traduttore è un lavoro preciso, che permette di ottenere l'eccellenza, oltre a un guadagno dignitoso. Si tratta di un'attività piacevole, appagante e, secondo alcuni partecipanti, anche divertente. Le parole prodotte e legate al post-editing, invece, fanno intuire l'esistenza di un divario tra i partecipanti: da una parte vi sono coloro che trovano in questa professione una grande utilità, oltre a un risparmio sui costi e sui tempi di traduzione; dall'altra, invece, vi sono coloro che si dimostrano totalmente avversi al post-editing, vedendolo come un'attività noiosa, imprecisa e sottopagata, volta allo sfruttamento dei traduttori. Tale discrepanza rappresenta l'insoddisfazione di una larga parte di traduttori nei confronti del post-editing.

#### **3.4.2 Analisi dei risultati**

In seguito al compito di associazione libera, sono state rivolte alcune domande a tutti i partecipanti. Innanzitutto, è stato chiesto se avessero mai sentito parlare di post-editing prima della compilazione di suddetto questionario (Figura 3). La maggioranza dei partecipanti, equivalente all'81.8% del totale, dichiara di sapere cosa sia il post-editing, e dunque, di averne già sentito parlare in precedenza. Risulta necessario sottolineare che tra questa porzione di partecipanti, la metà è costituita dai post-editor. Dunque, è possibile affermare con sicurezza che il 31.8% dei partecipanti rimanenti, facenti parte del pubblico generale e pur non essendo dei post-editor, sono a conoscenza dell'argomento trattato. Il 9.1% non lo conosce affatto e il restante 9.1% si dichiara incerto, sostenendo di non essere totalmente sicuro di averne sentito parlare. In generale, è possibile confermare che la maggioranza del pubblico generale conosce l'argomento.

Before this questionnaire, have you ever heard of post-editing?

132 responses

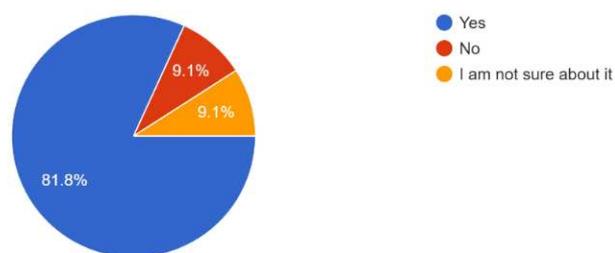


Figura 3: Conoscenza del post-editing prima del questionario

Particolarmente interessante risulta lo studio di come i partecipanti siano venuti a conoscenza di un determinato argomento e/o fenomeno, pertanto, dove ne abbiano sentito parlare (Figura 4). Non è stato imposto un limite di risposte, bensì ognuno dei partecipanti ha potuto selezionare più opzioni, se necessario. Il 53.4% dei partecipanti ha affermato di averne sentito parlare a lavoro: tale risposta è stata scelta prevalentemente dai post-editor e da alcuni traduttori, che, tuttavia, non si occupano di post-editing. La partecipante 101 (over 65) aggiunge, infatti, che il post-editing fa parte del proprio lavoro giornaliero e, avendo 30 anni di carriera alle spalle come traduttrice, ora lavora come direttrice di redazione nella propria azienda; la partecipante 32 (over 65), invece, ha affermato di averne sentito parlare presso la ÖVGD (Österreichischer Verband der allgemein beeideten und gerichtlich zertifizierten Dolmetscher), ovvero l'associazione austriaca degli interpreti giurati e giudiziari, presso la quale risulta probabilmente iscritta: la partecipante è infatti una traduttrice/interprete autonoma. Sono numerosi i partecipanti (34.5% del totale) che sono venuti a conoscenza del post-editing a scuola o all'università: si tratta soprattutto di studenti o di traduttori/post-editor, che hanno dunque appreso le basi di tale attività durante la propria formazione scolastica e universitaria. Internet e i social media rappresentano altri due canali tramite i quali i partecipanti sono venuti a conoscenza del post-editing: infatti le due opzioni sono state selezionate rispettivamente dal 47.4% e dal 31.9% dei partecipanti allo studio. Solamente due persone hanno dichiarato di averne sentito parlare alla televisione, 11 partecipanti hanno letto del post-editing all'interno di qualche libro, mentre 8 all'interno di una rivista o giornale. 19 partecipanti affermano di essere venuti a conoscenza del tema parlandone con la propria famiglia o con i propri amici. 3 traduttori e post-editor,

invece, avevano parlato con colleghi e clienti dell'argomento, arrivando, in alcuni casi, a intraprendere questa tipologia di professione. Workshop, forum relativi all'industria della traduzione e del post-editing, conferenze, gruppi professionali di discussione, webinar, rivolti specificatamente ai traduttori, rappresentano importanti mezzi di informazione e aggiornamento nell'ambito del post-editing, come riportato da 5 partecipanti. Solo una persona ne ha sentito parlare in un film. Infine, 3 partecipanti hanno affermato di non ricordare dove ne avessero sentito parlare (uno tra questi, infatti, alla precedente domanda aveva risposto di non essere sicuro di sapere cosa si intendesse con il termine "post-editing").

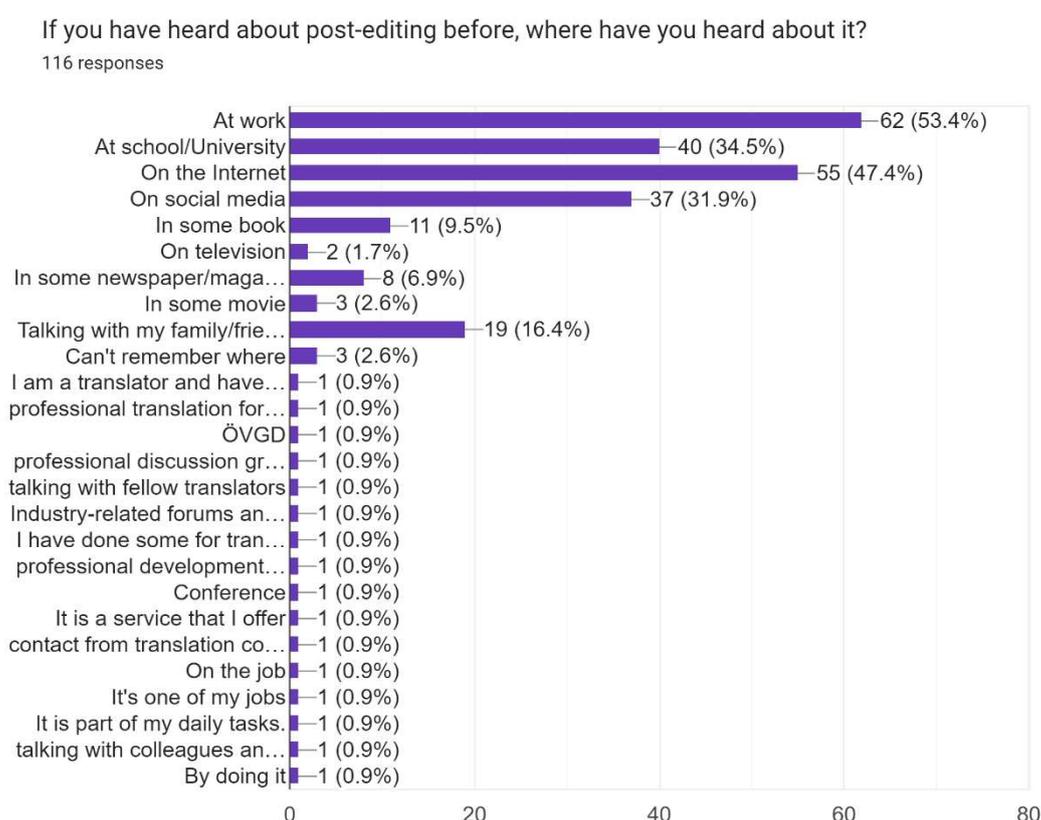


Figura 4: Dove si è sentito parlare di post-editing

È importante far presente che i partecipanti allo studio provengono da contesti lavorativi e educativi tra loro differenti, come riportato nella Figura 5: il 36.4% dei partecipanti, equivalente a 48 persone su 132 totali, dichiara di essere un lavoratore autonomo. Infatti, gran parte dei post-editor partecipanti si sono riconosciuti in questa categoria, in quanto freelance. Il secondo gruppo più ampio è rappresentato dagli studenti, corrispondente al 18.9% degli intervistati, ovvero 25 persone. 10 partecipanti

sono operai, 9 sono impiegati in ufficio (6.8% dei partecipanti), 8 sono insegnanti/maestri/professori (6.1% dei partecipanti), mentre 5 si dichiarano disoccupati, rappresentando il 3.8% del totale. Due partecipanti sono pensionati. Vi sono, poi, alcune tipologie di lavoro svolte solamente da uno o al massimo due partecipanti, quali tecnici, imprenditori, fisioterapisti, militari, farmacisti, sviluppatori di software e *transcreator*. Inoltre, alcuni traduttori e post-editor hanno voluto fornire delle indicazioni più specifiche riguardo alla propria occupazione: 3 si sono definiti come “traduttori/interpreti”; la partecipante 117 (45-54 anni) ha aggiunto che, oltre a essere una traduttrice freelance, possiede una piccola azienda, per la quale lavorano meno di 8 dipendenti; la partecipante 124 (55-64 anni) lavora sia come freelance, autonomamente, che come dipendente occasionale; la partecipante 101 (over 65) dice di aver lavorato per oltre 30 anni come traduttrice e di avere, ora, una propria azienda, in cui lavora come direttrice di redazione; infine, il partecipante 125 (45-54 anni) afferma di lavorare sia come traduttore che come project manager.

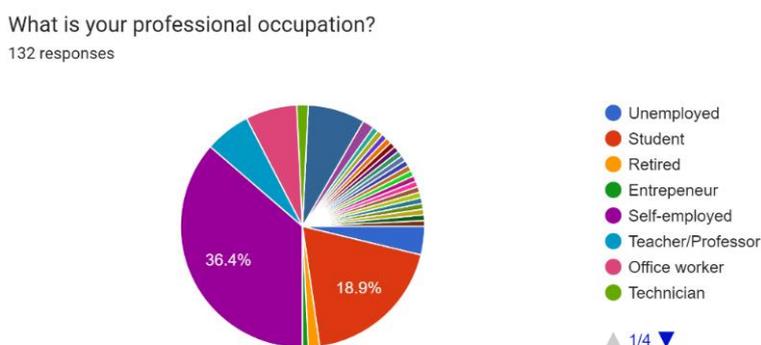


Figura 5: Occupazione professionale

Successivamente, è stato chiesto ai partecipanti se ritenessero che i metodi di remunerazione per il lavoro di post-editing e per quello di traduzione fossero equivalenti, ovvero se vengano usati gli stessi metodi di pagamento per entrambe le tipologie di lavoro (Figura 6). Il 60.6% dei partecipanti ha risposto di no, affermando che esistono metodi di remunerazione differenti per i due lavori. Tra coloro che hanno risposto “no” alla domanda in questione, vi sono, in maggioranza, i post-editor partecipanti allo studio e alcuni traduttori che, tuttavia, non si occupano di post-editing (53 persone su 80 ad aver risposto scelto l’opzione “no” sono post-editor), oltre a 13 studenti, probabilmente provenienti da corsi di laurea in discipline umanistiche o

orientati allo studio delle lingue straniere, alla traduzione e all'interpretariato. Più di un quarto dei soggetti che hanno preso parte allo studio (28.8% del totale), corrispondente a 38 persone su 132, hanno risposto "non lo so". Il questionario, infatti, è stato sottoposto a una gamma di persone provenienti da contesti lavorativi e educativi tra loro estremamente differenti e vari: pertanto, è comprensibile che vi sia una percentuale così alta di persone che non siano a conoscenza dei metodi di remunerazione normalmente utilizzati nel post-editing e nella traduzione. Infatti, tra questa categoria di partecipanti rientrano impiegati d'ufficio, alcuni studenti, operai, liberi professionisti e insegnanti, sia molto giovani (fascia di età tra i 18-24 anni) che meno (over 65). Tra questi è interessante la risposta della partecipante 32 (over 65), una traduttrice e interprete autonoma, che non ha idea di come venga pagato tale lavoro, non avendo mai lavorato nel campo del post-editing. Ciò significa che anche alcune persone che attualmente lavorano nel mondo della traduzione non sono correttamente a conoscenza di come venga retribuito il lavoro di post-editing. Infine, il restante 10.6% dei partecipanti, corrispondente a 14 persone, afferma che, secondo la propria opinione, i metodi di pagamento per i due lavori in questione sono uguali, scegliendo come risposta l'opzione "sì". Tra coloro che hanno risposto affermativamente alla domanda vi sono 12 traduttori freelance, autonomi, e post-editor, un paio dei quali ora pensionati. È vero, infatti, che in alcuni casi i metodi di pagamento per il post-editing e per la traduzione sono gli stessi. Ciononostante, vanno distinti, in quanto vengono utilizzati in maniera differente e hanno un peso diverso a seconda del lavoro di cui si sta parlando: il pagamento a parola, normalmente utilizzato nella traduzione e vantaggioso per tale impiego, non è sempre giudicato il migliore nel caso del post-editing, poiché, in alcune circostanze, si rivela meno equo e adatto. Proprio sull'analisi di queste differenze si basa la seconda parte del questionario, volta alla comprensione della percezione dei post-editor nei confronti dei diversi metodi di remunerazione del lavoro di post-editing.

Do you think the payment methods are the same for post-editing and translation?

132 responses

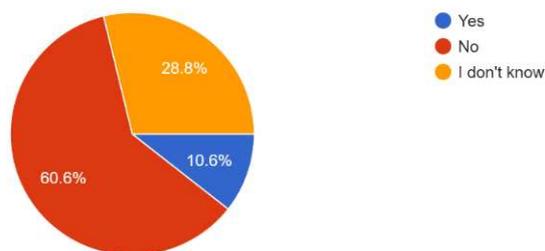


Figura 6: Confronto metodi di pagamento della traduzione e del post-editing

Successivamente è stato chiesto ai partecipanti se fossero a conoscenza dei metodi di pagamento esistenti per il post-editing (Figura 7). Tutti i partecipanti hanno fornito una risposta a questa domanda: il 50.8% del totale, corrispondente a 67 persone, hanno dichiarato di esserne a conoscenza. Tra coloro che hanno scelto la risposta “sì” rientrano i post-editor partecipanti (a eccezione di cinque), quattro studenti e due disoccupati (questi ultimi appartenenti a due fasce d’età, 18-24 anni e 25-34 anni: probabilmente sono dei neolaureati attualmente in cerca di lavoro), che hanno dunque affermato di essere a conoscenza dei diversi metodi di pagamento. Il 32.6% dei partecipanti, ovvero 43 persone, hanno invece detto di non sapere quali siano i metodi di pagamento esistenti per il post-editing: tale opzione è stata scelta in particolare da coloro che alla precedente domanda avevano risposto “non lo so”: si tratta, quindi, di soggetti provenienti da ambiti di studio o lavorativi lontani dalla traduzione e dal post-editing. Alcuni avevano dichiarato precedentemente di non aver mai sentito parlare di post-editing, o perlomeno non erano certi di averne sentito parlare. Infine, il 16.7% dei soggetti partecipanti, pari a 22 persone, hanno dichiarato di non essere certi di essere a conoscenza dei diversi metodi di pagamento esistenti per il post-editing: anche in questo caso si tratta di alcuni partecipanti, appartenenti al pubblico generale, che alla precedente domanda avevano risposto con “non lo so”, oltre a cinque post-editor stessi (che non avevano ben capito cosa si intendesse con “metodo di pagamento”). Le risposte raccolte permettono di intuire che, all’interno del pubblico generale e, dunque, anche tra i potenziali clienti, vi è ancora molta incertezza riguardo al pagamento del post-editing.

Are you aware of the existing payment methods for post-editing?

132 responses

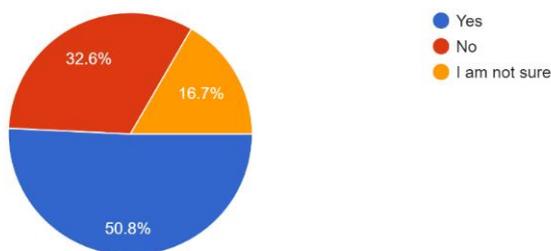


Figura 7: Conoscenza dei metodi di pagamento del post-editing

Per ottenere una visione più precisa, è stato richiesto ai soggetti partecipanti al questionario di elencare i metodi di pagamento da loro conosciuti. Sebbene 78 partecipanti abbiano fornito una risposta al quesito in questione, non tutti i risultati ottenuti sono stati presi in considerazione, in quanto giudicati non particolarmente interessanti ai fini della ricerca. Le risposte considerate valide sono riportate nella Figura 8.

I metodi di remunerazione maggiormente nominati sono il pagamento a parola (citato da 33 persone) e quello su base oraria (citato da 32 persone). Tra i partecipanti, 8 non hanno indicato un metodo di pagamento preciso, ma si sono limitati a sottolineare che il post-editing viene pagato decisamente meno rispetto alla traduzione: altri 7 traduttori, infatti, affermano che normalmente il post-editing viene pagato circa la metà della traduzione, mentre un partecipante sostiene che sia, in alcuni casi, pari al 30% della traduzione. Le percentuali, dunque, variano a seconda della propria esperienza personale; tuttavia, si può affermare che, come riportato da 4 post-editor, la retribuzione si aggira normalmente tra il 50% e il 70% del prezzo previsto per il lavoro di traduzione. Il metodo retribuzione basato sul progetto o sul lavoro richiesto dal cliente è stato nominato da 7 partecipanti, mentre sono stati altrettanti coloro che hanno menzionato il pagamento basato su ogni parola del testo sorgente. Tra i post-editor, inoltre, 3 hanno nominato il pagamento a parola, ma ridotto (riferito al pagamento classico utilizzato nella traduzione), mentre 2 partecipanti hanno affermato che esiste un pagamento ridotto per i *fuzzy match*. La retribuzione in base al numero di linee è stata citata da 5 traduttori, come anche il pagamento basato su ogni segmento cambiato o modificato. Altri metodi di remunerazione nominati dai partecipanti sono stati il

pagamento per pagina/pagina standard (4 persone) e quello basato sulla distanza di modifica (2 persone). Il calcolo per la remunerazione del post-editing, secondo uno dei traduttori, è composto da una frazione del pagamento esistente per la traduzione (*from scratch*), basata sul tempo idealmente necessario per la modifica di ogni parola. Un paio di partecipanti si sono limitati a sottolineare che il pagamento del post-editing sia troppo basso (“poorly paid”), mentre un altro sostiene che i metodi utilizzati nel post-editing siano gli stessi di cui ci si serve normalmente per la remunerazione della traduzione svolta con strumenti CAT, ma con dei pagamenti notevolmente più bassi: tale affermazione è sostenuta anche da un altro traduttore, il quale aggiunge che, in caso di *full match*, non vi sia alcun pagamento. Tra i meno menzionati, citati da solo un partecipante, vi sono il pagamento calcolato sul tempo speso per il post-editing e quello basato sulla performance. Un post-editor, invece, afferma che la retribuzione del post-editing è in pari quella normalmente prevista per la revisione. Infine, un partecipante ha nominato il TER, “Translation Edit Rate”, spiegando in cosa consiste tale sistema di remunerazione: il TER, infatti, misura, in percentuale, la quantità di cambiamenti apportati da un post-edito a un testo, precedentemente tradotto tramite l’uso di un sistema di traduzione automatica. Dunque, il pagamento è basato sulla quantità di cambiamenti e sul tempo effettivamente impiegato per lo svolgimento di tale lavoro. È possibile affermare con certezza che i metodi di pagamento del post-editing, conosciuti dai post-editor e dal pubblico generale, sono molteplici; inoltre, ognuno dei partecipanti ha fornito un ulteriore contributo, esprimendo la propria opinione in base all’esperienza personale. Ciononostante, è appurato che i metodi più conosciuti, in quanto maggiormente citati, sono quello a parola e quello su base oraria. La presenza di una vasta quantità di sistemi di remunerazione differenti, tuttavia, segnala la mancanza di indicazioni specifiche e univoche sul tema: proprio per questa ragione, sussistono delle situazioni economicamente svantaggiose per alcuni post-editor, i quali non ricevono una corretta compensazione per il loro lavoro.

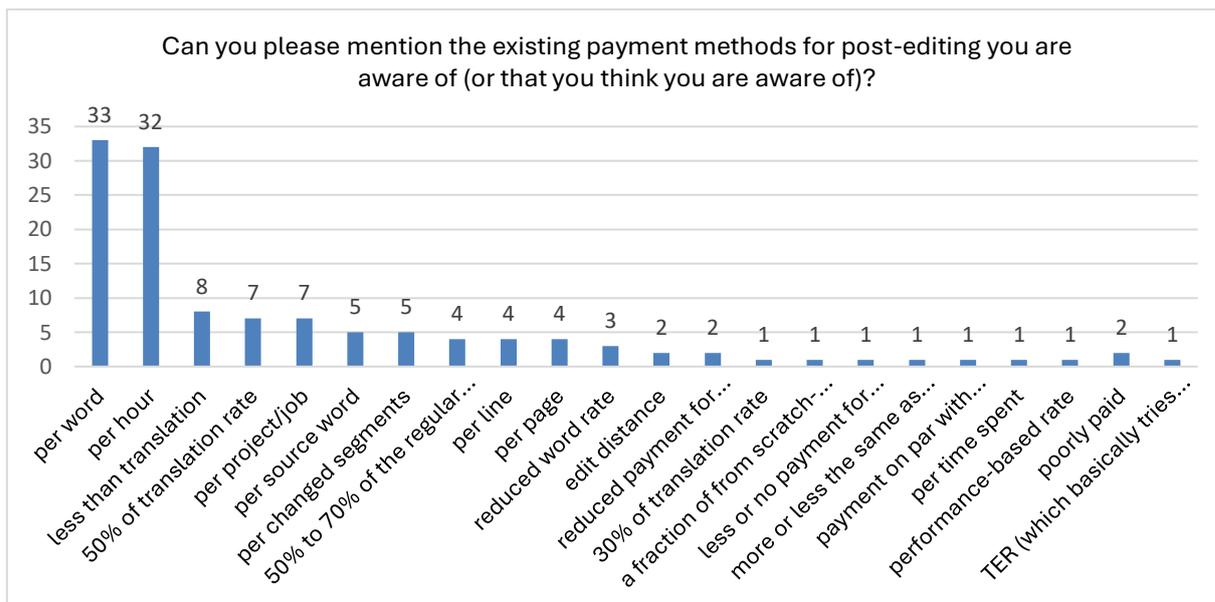


Figura 8: Metodi di pagamento del post-editing di cui si è a conoscenza

### 3.5 Seconda sezione: il punto di vista dei post-editor

La seconda parte del questionario è stata studiata per essere sottoposta esclusivamente ai traduttori che si occupano di post-editing, o che se ne sono occupati nel passato. Le domande, infatti, sono volte a comprendere il punto di vista dei post-editor sulla remunerazione del proprio lavoro, evidenziando le principali problematiche, al fine di ottenere una visione più chiara su quale sia, secondo la maggioranza, il metodo di pagamento considerato più equo. Nella seguente sezione sono state proposte ai partecipanti una serie domande chiuse, a risposta singola o multipla. Come precedentemente riportato, la metà dei partecipanti (66 persone) rientra all'interno della categoria dei post-editor.

#### 3.5.1 Analisi dei risultati

Il post-editing, al giorno d'oggi, è un lavoro estremamente richiesto. Pertanto, è stato reputato interessante chiedere ai post-editor le motivazioni che li hanno spinti a intraprendere tale professione. Trattandosi di una domanda a risposta multipla, diversi partecipanti hanno selezionato più di un'opzione come possibile risposta (Figura 9). La maggioranza degli intervistati, pari al 74.6% del totale dei post-editor, ha dichiarato che il post-editing è, per l'appunto, giudicato essenziale all'interno del mercato della traduzione, divenendo enormemente richiesto dai clienti: per diversi post-editor, dunque, è una scelta obbligata, dettata dalle richieste del mercato target, che lascia meno spazio alla traduzione vera e propria, come riportato dalla partecipante 18 (45-54 anni) e

dalla partecipante 25 (55-64 anni). In alcuni casi il post-editing è l'unica opzione per ottenere un determinato lavoro (“It was decided by the translation agency, I wouldn't get the job<sup>16</sup>” – partecipante 48, donna, 25-34 anni). Altri 5 partecipanti hanno sottolineato che si è trattata di una richiesta da parte dei clienti, che spesso ritengono di risparmiare sui tempi e sui costi in questa maniera. In realtà, come sottolineato dalla partecipante 117 (45-54 anni) l'output della MT è più scadente di una traduzione umana, e contiene, oltretutto, errori e imprecisioni linguistiche. L'impossibilità della scelta è, tra l'altro, uno dei fattori che contribuiscono a rendere il post-editing un lavoro poco apprezzato dai traduttori. Il 22.4% dei partecipanti ha sottolineato, tuttavia, l'esistenza di un'utilità di base del post-editing, considerandolo, infatti, una buona possibilità per aumentare la propria produttività. In effetti, nel caso in cui l'output della MT presenti una buona qualità, non è necessario apportare numerose modifiche, svolgendo il lavoro più velocemente e ottenendo, di conseguenza, un aumento della produttività. Il post-editing, secondo il 22.4% dei partecipanti, permette di risparmiare molto tempo: anche in questo caso l'affermazione è da considerarsi valida solamente nelle circostanze in cui la qualità dell'output sia buona. In caso contrario, il post-editing si rivela particolarmente dispendioso in termini di tempo. Il 19.4% dei post-editor, inoltre, ha dichiarato di aver intrapreso tale professione al fine di incrementare i propri profitti. Alcuni partecipanti, 4 per la precisione, hanno affermato di aver intrapreso questa professione per curiosità: tra questi, il partecipante 67 (45-54 anni) ha dichiarato di apprezzare il lavoro di post-editing e di svolgerlo volentieri, mentre il partecipante 37 (over 65) ha espresso un parere negativo a riguardo, affermando che non ne valga la pena. Solamente 3 post-editor, ovvero il 4.5% dei partecipanti, hanno dichiarato occuparsi di post-editing per l'esistenza di un metodo di remunerazione equo, opzione che non è stata presa in considerazione dai restanti partecipanti.

---

<sup>16</sup> “È stato deciso dall'agenzia di traduzione. Non avrei ottenuto il lavoro” (tradotto da me).

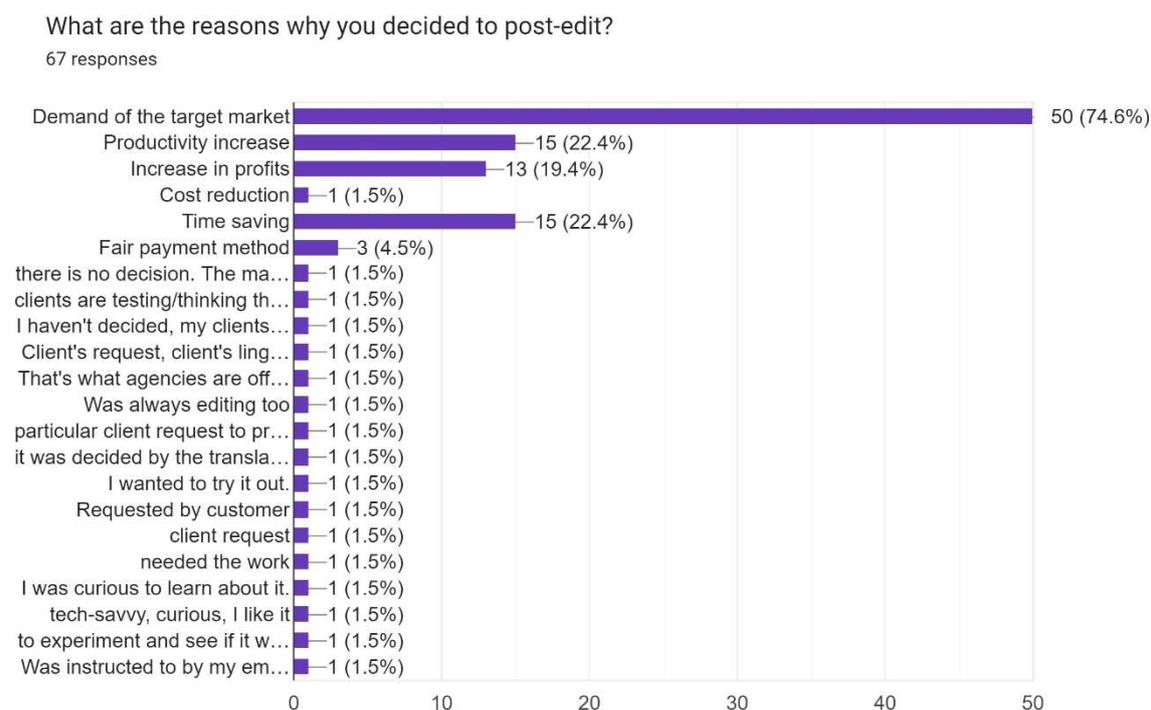


Figura 9: Motivazioni per cui si è scelto di effettuare il post-editing

Il pubblico generale e i clienti non sono generalmente ben informati sul post-editing e sulla sua retribuzione. Vi è una scorretta percezione del lavoro, oltre a una tendenza a sottovalutare le difficoltà che i post-editor possono incontrare nel suo svolgimento. Una delle domande proposte ai partecipanti, difatti, ha visto come oggetto la percezione del post-editing da parte del pubblico generale (Figura 10): secondo l'opinione di 58 post-editor su 66, equivalente all'88% del totale, il pubblico generale non ha una corretta percezione sul tema. Tale disinformazione provoca degli effetti negativi sul lavoro di post-editing, in quanto all'interno della vasta categoria "pubblico generale" rientrano anche dei potenziali clienti. Una scorretta percezione del tema e di quale possa essere un sistema di pagamento equo, comporta l'adozione di un metodo di compensazione inadeguato. In effetti, nessuno dei partecipanti ha affermato che il pubblico generale possiede una visione e percezione appropriata del tema. I restanti 8 partecipanti, corrispondenti al 12% del totale, si dichiarano indecisi, affermando di non sapere con certezza se il pubblico abbia una giusta percezione o meno.

Do you think the general public has a correct perception of post-editing and its compensation in terms of money?



Figura 10: Percezione del pubblico generale sul post-editing

La successiva domanda ha come oggetto la remunerazione dei post-editor, ovvero come vengono normalmente pagati per lo svolgimento di tale lavoro (Figura 11): il 66.2% dei partecipanti, corrispondente a 43 post-editor, dichiara di essere normalmente retribuito tramite il pagamento a parola. Questa metodologia, infatti, è considerata la più classica e standard, sebbene non sempre sia in grado di soddisfare pienamente le esigenze dei post-editor. Il 18.5% dei partecipanti, corrispondente a 12 persone, invece, riceve normalmente una retribuzione su base oraria. Un post-editor dichiara di essere normalmente pagato secondo la distanza di modifica, mentre un altro a progetto. Alcuni post-editor hanno fornito una risposta più precisa a riguardo o diversa dalle opzioni già presenti, riportando la propria esperienza: sono stati elencati, come altri metodi di remunerazione, quello per pagina standard, il pagamento per *weighted word* (parola pesata); un paio di post-editor affermano di essere pagati a volte su base oraria e a volte a parola, a seconda di quale venga concordato con il cliente; la partecipante 75 aggiunge, inoltre, di non accettare lavori la cui compensazione è troppo bassa. Il partecipante 37 afferma di non svolgere più lavori di post-editing, come il partecipante 65, il quale dichiara di aver ricevuto, nel precedente impiego lavorativo, lo stesso salario sia per i lavori di post-editing sia per quelli di traduzione o altri lavori di ufficio.

Sembra, dunque, che i clienti siano maggiormente predisposti nei confronti di una retribuzione a parola. Da alcune risposte si può dedurre che vi sia un grado di insoddisfazione da parte di alcuni partecipanti: successivamente, infatti, verranno proposte le analisi dei dati relativi al grado di soddisfazione dei post-editor nello svolgimento del proprio lavoro.

How do you usually get paid when you do post-editing?  
65 responses



Figura 11: Metodo di pagamento usuale per il post-editing

Una porzione consistente di post-editor non è soddisfatta, tuttavia, del metodo di remunerazione normalmente utilizzato dai propri clienti (Figura 1): infatti, 32 traduttori su 65 hanno dichiarato di non sentirsi appagati dal metodo di retribuzione ricevuto, rappresentando il 49.2% dei post-editor partecipanti allo studio. Circa la metà dei traduttori che si occupano di post-editing, pertanto, ritengono migliore un metodo di pagamento diverso da quello normalmente ricevuto. Tra i 32 post-editor che si definiscono scontenti rientrano 24 traduttori che normalmente vengono pagati a parola, rappresentando il 75% dei partecipanti insoddisfatti. Si reputano soddisfatti del metodo di pagamento ricevuto, invece, 33 post-editor su 65, ovvero il 50.8% del totale. Anche in questo caso vi sono molti post-editor che normalmente ricevono una remunerazione a parola (21 persone) o su base oraria (10 persone). La retribuzione a parola, dunque, non è considerata in tutti i casi un'opzione da scartare, in quanto dipende da come viene calcolata.

Are you satisfied with the payment method by which you get paid?  
65 responses

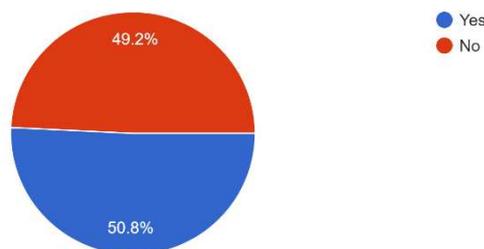


Figura 12: Soddisfazione del metodo di pagamento normalmente ricevuto

È stato quindi chiesto ai singoli post-editor partecipanti allo studio in questione quale sia, secondo la propria opinione, il metodo di pagamento più adatto a tale lavoro (Figura

13). Le risposte ottenute sono state numerose e tra loro varie. 2 traduttori giudicano migliore il metodo basato sulla distanza di modifica: entrambi vengono normalmente pagati a parola, metodo di cui non sono soddisfatti. 30 post-editor ritengono che il metodo di retribuzione più equo sia quello su base oraria, rappresentando il 44.1% del totale: si tratta soprattutto di post-editor normalmente pagati a parola (16), secondo la distanza di modifica (1), a progetto (1) o con un altro metodo non precedentemente specificato (1), insoddisfatti del pagamento ricevuto, e di post-editor di norma pagati su base oraria (11) e soddisfatti di tale metodo retributivo. Sono 3 i partecipanti che ritengono il pagamento calcolato in base al progetto la metodologia più equa: i partecipanti 52 e 62 affermano che, pur essendo soddisfatti del pagamento a parola, quello calcolato in base al progetto sia il metodo più equo; lo stesso vale per il partecipante 95, pagato di volta in volta diversamente. 22 post-editor, corrispondenti al 33.8% del totale, invece, ritengono la remunerazione a parola quella più equa: si tratta principalmente di traduttori che attualmente ricevono un pagamento a parola e ne sono soddisfatti (18), o post-editor che ricevono una retribuzione su base oraria, ma non la ritengono abbastanza adeguata (2), oltre a due partecipanti che non si occupano più di post-editing. Il partecipante 82 dichiara che il metodo più equo sia per pagina standard, con cui attualmente viene remunerato. Una sola post-editor, normalmente pagata a parola e insoddisfatta di tale sistema di remunerazione, sostiene che il post-editing dovrebbe essere pagato come la traduzione. Due post-editor non si sono espressi chiaramente, ovvero non hanno indicato un metodo di pagamento, ma, secondo la propria opinione, innanzitutto sarebbe necessario tenere conto della qualità dell'output della traduzione prodotta dalla MT, del progetto e della coppia linguistica, elementi fondamentali all'individuazione della forma di retribuzione maggiormente adeguata. Due partecipanti, invece, dichiarano più equo un approccio combinato, che prevede l'utilizzo sia del pagamento su base oraria sia quello basato sulla distanza di modifica, calcolato, tuttavia, con criteri che differiscono da quelli normalmente utilizzati negli strumenti CAT (partecipante 28). Tuttavia, se la distanza di modifica risulta troppo alta, sarebbe opportuno optare per un pagamento totalmente su base oraria o, addirittura, ritradurre l'intero testo, applicando i normali criteri utilizzati per il pagamento della traduzione (partecipante 18). Quest'ultima opzione è sostenuta anche dalla partecipante 124, la quale dichiara che, quando l'output presenta una qualità molto scarsa, la

soluzione migliore sarebbe quella di riscrivere la traduzione dal principio: in tale caso, il pagamento a parola per l'intera traduzione sarebbe il più adatto e vantaggioso. Secondo uno dei partecipanti, la retribuzione per il post-editing dovrebbe essere formata dall'insieme di due tipologie di remunerazione: una base costituita dal pagamento a parola, a cui, tuttavia, dovrebbe essere aggiunto un extra, una retribuzione aggiuntiva, in base al livello di editing e/o di ricerca richiesto. Infine, una traduttrice (partecipante 77), dichiara di non poter affermare con sicurezza quale sia il metodo di pagamento maggiormente equo in quanto non ha ancora avuto l'occasione di sperimentare diversi metodi di remunerazione, poiché normalmente viene pagata a parola.

According to your opinion, what is the most suitable payment model when post-editing?

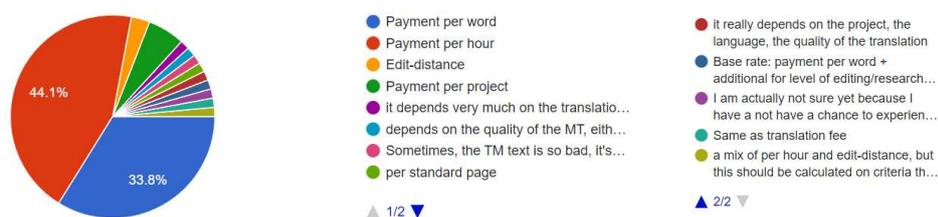


Figura 13: Metodo di pagamento migliore per il post-editing

È stato inoltre chiesto ai post-editor di esprimere il proprio grado di soddisfazione riguardo al guadagno ottenuto con il post-editing, se messo a confronto con lo sforzo cognitivo richiesto da tale lavoro (Figura 14). Tutti i post-editor che hanno preso parte allo studio hanno fornito una risposta a tale domanda. La maggioranza di essi, corrispondente ovvero a 38 persone su un totale di 66 ed equivalente al 57.6% dei partecipanti, ha definito il rapporto “guadagno ottenuto/sforzo impiegato” nel post-editing non sufficientemente adeguato: si può evincere, dunque, che l'impegno richiesto dal lavoro di post-editing non è limitato. Lo sforzo cognitivo, infatti, è alle volte superiore a quello normalmente richiesto dalla revisione o dalla traduzione. Ciò accade poiché i sistemi di traduzione automatica neurali sono in grado di produrre un output apparentemente perfetto, che, tuttavia, nasconde degli errori. Il post-editor, dunque, deve prestare molta attenzione all'output al fine di riuscire a individuare i diversi errori presenti. Il 27.3% dei partecipanti ha affermato che i pagamenti ricevuti sono adeguati allo sforzo richiesto mentre il restante 15.2% ha scelto la risposta “non lo so”, dimostrandosi incerto sulla questione: l'impegno previsto dal post-editing è strettamente

legato alla qualità dell'output prodotto dal sistema di traduzione automatica, che a sua volta dipende da una serie di variabili quali il contenuto del testo di partenza (dunque, la tipologia di testo) e la coppia linguistica presa in considerazione. Pertanto, lo sforzo richiesto potrebbe variare di volta in volta, in base alle circostanze sopracitate, non permettendo di fornire una risposta univoca alla domanda in questione, secondo alcuni dei partecipanti.

Are rates adequate for the effort required for PE? Think about what you can earn in a day only doing PE.

66 responses

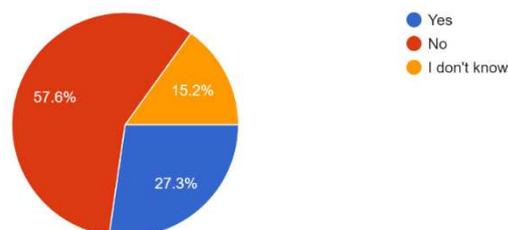


Figura 14: Adeguatezza dei pagamenti per il post-editing in relazione allo sforzo richiesto

È stato chiesto ai post-editor, poi, di immaginare di passare un'intera giornata lavorativa effettuando solamente post-editing e di paragonare gli ipotetici guadagni ottenuti nel corso di tale giornata con quelli normalmente conseguiti durante un giorno lavorativo, speso solamente a tradurre. Lo scopo della domanda in questione è quello di comprendere se il rapporto tra il guadagno ottenuto tramite il lavoro di post-editing è adeguato al tempo impiegato per il suo svolgimento, se messo in relazione con i normali ricavi ottenuti con la traduzione; dunque, si vuole constatare se i profitti conseguiti con le due tipologie di lavoro prese in esame sono tra loro simili. I risultati (riportati nella Figura 15) mostrano che quasi la metà dei partecipanti che hanno risposto a tale domanda, ovvero il 47.7% dei post-editor, dichiara che i propri guadagni, ottenuti con il lavoro di post-editing sono inferiori rispetto a quelli ottenuti con la traduzione. Infatti, solo il 23.1% dei partecipanti, ovvero 15 su 65 post-editor, afferma di guadagnare di più con il lavoro di post-editing che con quello di traduzione. Infine, una porzione significativa di partecipanti, corrispondente al 29.2% dei post-editor, non riscontra una differenza tra i guadagni derivanti dalla traduzione e quelli derivanti dal post-editing, che risultano, dunque, uguali o molto simili. In realtà, l'ipotesi di passare un'intera giornata solamente occupandosi di post-editing viene giudicata da diversi post-editor e

traduttori come remota, in quanto lo sforzo cognitivo richiesto è normalmente troppo elevato. Ciò significa che, anche se il post-editing permettesse di guadagnare di più, sarebbe comunque improbabile poterlo svolgere come unico lavoro.

In your opinion, how fair are post-editing rates compared to the time spent?  
65 responses

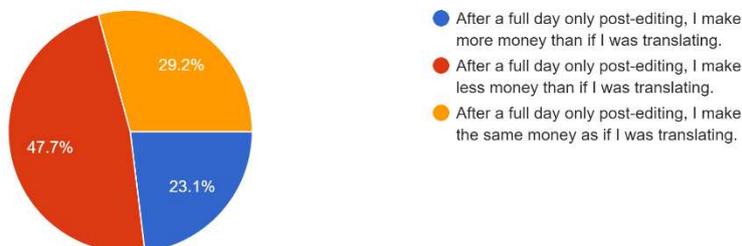


Figura 15: Adeguatezza del pagamento per il post-editing in relazione al tempo speso

Nell'ultima serie di domande è stato richiesto ai post-editor di esprimere la propria posizione rispetto a sei affermazioni proposte. Le risposte ottenute sono riassunte dalla Figura 16. Nella prima affermazione viene dichiarato che il pagamento a parola è il metodo di retribuzione più equo per il post-editing: riguardo a tale affermazione, 4 post-editor hanno scelto la risposta "non lo so", senza dunque esprimersi chiaramente a riguardo; 19 affermano di essere parzialmente d'accordo mentre 15 sostengono di essere parzialmente contrari a tale affermazione; 12 sono fortemente convinti che l'affermazione in questione sia corretta mentre altri 15 sono altamente contrari a essa. Ciò significa che 30 partecipanti ritengono che il pagamento a parola non sia il metodo remunerativo più equo tra i presenti, mentre 31 si sono espressi in maniera parzialmente e altamente positiva sul tema. Si può quindi dire con certezza che circa la metà dei partecipanti si dimostra avverso al pagamento a parola, che, tuttavia, è spesso il metodo di pagamento più utilizzato dai clienti per la retribuzione dei post-editor.

Nella successiva dichiarazione, al contrario, viene chiesto ai partecipanti di esprimere la propria opinione nei confronti della seguente affermazione: "Il pagamento su base oraria è il metodo di pagamento più equo". Tra i post-editor, 3 sostengono di non sapere come esprimersi a riguardo e che posizione prendere; 19 traduttori si definiscono "parzialmente d'accordo" mentre 25 sono fortemente d'accordo; nove sono parzialmente in disaccordo con l'affermazione e infine otto traduttori sono fortemente contrari. Ciò significa che 44 post-editor dimostrano di favorire, parzialmente e/o

totalmente, la retribuzione su base oraria, rappresentando una cifra superiore rispetto a quella rappresentante i sostenitori della retribuzione a parola. Sono, infatti, solo 17 coloro che si sono espressi negativamente nei confronti della remunerazione su base oraria. Pertanto, vi è un'ulteriore conferma che il pagamento calcolato su una base oraria potrebbe rappresentare un metodo di retribuzione equo, adatto, e apprezzato da più della maggioranza dei traduttori che si occupano di post-editing.

La terza affermazione ha come oggetto i guadagni ottenuti dal post-editing e il proprio grado di soddisfazione (“Sono soddisfatto/a dei guadagni ottenuti dal post-editing”): 25 post-editor si definiscono parzialmente soddisfatti (15 hanno risposto “abbastanza d'accordo”) o totalmente soddisfatti (10 si sono definiti “totalmente d'accordo”) dei propri guadagni, rappresentando meno della metà dei partecipanti che hanno risposto al quesito. Infatti, 34 traduttori hanno affermato di essere abbastanza insoddisfatti (12 persone) o totalmente insoddisfatti (22 persone). 5 post-editor hanno scelto l'opzione “non lo so”, senza, dunque, esprimersi riguardo al tema. È possibile affermare, perciò, che si nota un malcontento generale tra i post-editor, relativo ai propri guadagni, spesso troppo bassi e non adeguati allo sforzo richiesto dal lavoro. Tale insoddisfazione è ulteriormente aggravata dall'utilizzo di un metodo di remunerazione inadatto: l'applicazione di un sistema retributivo equo permette di tenere in considerazione le diverse variabili ed esigenze, portando, conseguentemente, all'aumento dei guadagni. In effetti, 23 dei 34 post-editor che hanno dichiarato di non essere soddisfatti dei propri guadagni per il post-editing sono normalmente retribuiti a parola: ciò dimostra nuovamente che, in molti casi, il metodo di remunerazione a parola è inadeguato e non permette di ottenere un guadagno sufficiente.

In seguito, è stato chiesto ai post-editor di esprimersi in merito alla seguente dichiarazione: “Il post-editing riduce i miei costi”. 11 traduttori hanno dichiarato di non sapere come esprimersi a riguardo; sono stati 10 i post-editor che si sono dichiarati abbastanza d'accordo con l'affermazione e 11 coloro che si sono ritenuti parzialmente in disaccordo; 5 post-editor si sono espressi come fortemente d'accordo con la dichiarazione in questione mentre 25 sono molto in disaccordo. Questo significa che per 38 partecipanti, che costituiscono la maggioranza delle risposte ottenute, il lavoro di post-editing non diminuisce totalmente o parzialmente i costi da sostenere. Dunque, oltre a non essere soddisfatti dei propri guadagni, probabilmente troppo bassi se

comparati alle aspettative di partenza o ai guadagni normalmente ottenuti con il lavoro di traduzione o di revisione, molti traduttori che si occupano di post-editing non riscontrano, normalmente, una riduzione nei costi.

La successiva affermazione è di carattere provocatorio: “Il post-editing è una perdita di tempo e denaro”. Oltre a 6 traduttori che hanno scelto l’opzione “non lo so” e non hanno espresso un chiaro giudizio, la maggioranza dei post-editor che hanno risposto al quesito, ovvero 46 partecipanti, si sono dimostrati parzialmente disaccordo (27 persone) o totalmente disaccordo (19 persone). Infatti, sebbene il post-editing non sia un’attività generalmente molto amata dai traduttori, che normalmente preferiscono concentrarsi sulla traduzione stessa, poiché ritenuta più creativa e stimolante, sono pochi i post-editor che definiscono tale lavoro una perdita di tempo e di denaro: viene dimostrato, quindi, che il post-editing è comunque un’attività che permette di ottenere un profitto, sebbene spesso basso, e che, in condizioni ottimali (come, ad esempio, una buona qualità dell’output prodotto dalla MT), permette di aumentare la propria produttività e i propri guadagni. Tuttavia, sono 14 i post-editor che ritengono, parzialmente (12 persone) o totalmente (2 persone), che tale lavoro sia poco proficuo, soprattutto se comparato al tempo che richiede. Si nota, dunque, l’insoddisfazione di una piccola porzione di post-editor che, sebbene sia limitata, è da considerare comunque significativa. Alcuni traduttori, infatti, si occupano di post-editing in seguito a un aumento della domanda di tale lavoro all’interno del mercato della traduzione: il post-editing rappresenta, in alcuni casi, una scelta obbligata, risultando un compito da svolgere malvolentieri.

Infine, l’ultima affermazione ha come obiettivo principale quello di comprendere il punto di vista dei post-editor in merito al rapporto tra tempo speso per il post-editing e il rispettivo guadagno ricevuto. Lo scopo è quello di comprendere se tale rapporto sia adeguato o meno. Anche in questo caso 5 partecipanti hanno scelto l’opzione “non lo so”, senza esprimersi chiaramente. Tra i restanti post-editor, sono 34 coloro che ritengono di non essere adeguatamente pagati per il tempo impiegato nell’effettuare il post-editing: in particolare 11 si sono dichiarati parzialmente in disaccordo con l’affermazione e 23 totalmente contrari. Anche in questo caso il grado di insoddisfazione espresso dai partecipanti è relativamente alto. Una possibile causa di tale insoddisfazione potrebbe risiedere nel fatto che, come già sottolineato, molto spesso

i metodi di pagamento proposti e utilizzati dai clienti non tengono conto del tempo necessario per la lettura del testo, la sua comprensione, eventuali ricerche aggiuntive sul testo/argomento, l'individuazione degli errori, la rielaborazione, se necessario, dei segmenti e la rilettura del lavoro finale. Esaminando i risultati ancor più nel dettaglio, tra gli insoddisfatti rientrano soprattutto i post-editor che normalmente ricevono un pagamento a parola, rappresentando 21 persone su 34. 26 post-editor, invece, si dichiarano parzialmente (16 persone) e totalmente (10 persone) soddisfatti del rapporto tra il tempo speso per il post-editing e il suo relativo guadagno.

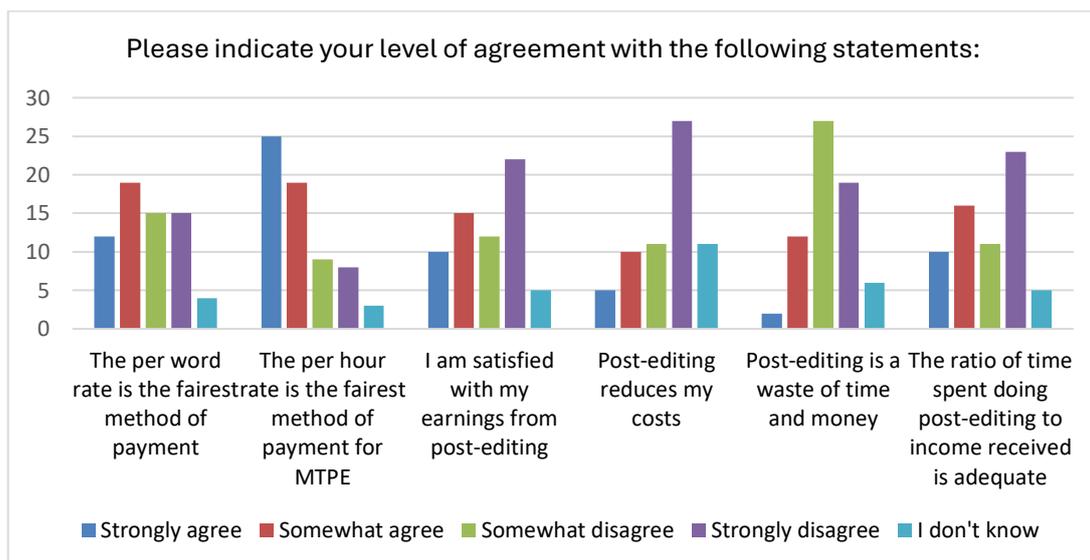


Figura 16: Grafico riassuntivo delle opinioni espresse

### 3.5.2 Considerazioni finali fornite dai post-editor

A conclusione dell'analisi dei dati raccolti, vengono riportati alcuni degli ulteriori contributi apportati da diversi post-editor nella sezione finale, dedicata a commenti aggiuntivi, in cui hanno descritto alcune esperienze personali in merito al lavoro di post-editing. Anche in questo caso sono stati scelti gli interventi giudicati maggiormente significativi e approfonditi, che forniscono degli spunti di riflessione particolarmente interessanti. Ad esempio, la partecipante 17 (46-54 anni, dunque con molta probabilità una traduttrice esperta), afferma come di seguito riportato:

Personalmente, ritengo che la tariffa oraria sia l'unica modalità accettabile negli incarichi di MTPE perché consente di dar conto dell'effettivo sforzo (tempo) richiesto dagli specifici progetti di post-editing. [...] Così come i CAT Tool, la traduzione automatica è un semplice strumento applicabile ai processi di traduzione che, se impiegato in modo appropriato, potrebbe aumentare l'efficienza e la qualità delle traduzioni. A mio avviso, l'appropriatezza nell'uso di questo strumento può sintetizzarsi in 1. applicazione della MT solo a determinate tipologie testuali; 2. gestione del documento fin dalla fase di redazione in funzione dell'uso della MT (applicazione dei parametri del linguaggio naturale

controllato); 3. corretta gestione del processo (redazione/pre-editing/MT/post-editing/revisione); 4. affidamento degli incarichi di MTPE a professionisti specializzati. Il problema è che la MT è spesso intesa meramente come un modo per tagliare i costi nella gestione della documentazione multilingue. Non è intesa come uno strumento, ma come un sistema in grado di sostituire (per ora in parte) la traduzione "umana". Viene applicata indiscriminatamente a qualsiasi tipologia testuale e contenuto su testi che non sono redatti in base ai parametri del linguaggio naturale controllato, né sono sottoposti a pre-editing prima del passaggio in MT. Il risultato è una qualità dell'output di MT scadente e testi pre-tradotti che spesso necessitano di una vera e propria ritraduzione. In generale, inoltre, non si considera il fatto che, in tutte le fasi, i processi di MT richiedono competenze specifiche in parte diverse da quelle richieste nella redazione, gestione e traduzione dei documenti senza l'uso di strumenti di automazione. Ipotizzando che il committente si aspetti a fine processo una qualità del testo pari a quella di una traduzione eseguita da un traduttore competente, l'applicazione di una tariffa a parola del testo sorgente (quindi una tariffa fissa, che non tiene conto della qualità dell'output di MT) consente al committente stesso di esternalizzare i costi non sostenuti per la preparazione dei testi sul post-editor, per il quale la qualità dell'output di MT incide sul tempo necessario a portare a termine l'attività di MTPE. Questa situazione non incentiva i committenti a migliorare la gestione del processo nel suo complesso. Al contrario, negli anni ho notato una situazione paradossale: a fronte del notevole miglioramento delle prestazioni dei sistemi di MT, la qualità dei progetti di MTPE è drasticamente diminuita. (Partecipante 17, post-editor, donna, 45-54 anni)

La partecipante in questione, pertanto, ribadisce l'importanza di un pagamento su base oraria, giudicato da lei come unico mezzo di pagamento possibile, in quanto l'unico in grado di tenere conto delle variabili che compongono il processo di post-editing. Inoltre, evidenzia ulteriormente la scarsa preparazione sull'argomento da parte dei clienti e le pretese di coloro che commissionano lavori di post-editing, che vorrebbero un risultato di ottima qualità, pari a quella umana, ma con i prezzi del post-editing e in tempi brevi. Non è raro che, come affermato dalla partecipante 17, i post-editor debbano riscrivere intere frasi, perché la qualità è così bassa che una nuova compilazione risulta più veloce della modifica stessa. Viene ribadita la scarsa attenzione al tipo di contenuto: la traduzione automatica, infatti, non è adatta a qualsiasi tipo di testo, sebbene i clienti non ne tengano conto. Il loro principale obiettivo risiede nell'abbassamento dei costi, e puntano solamente a questo.

La partecipante 68 (post-editor, donna, 45-54 anni) afferma che non è possibile lavorare un giorno intero occupandosi solo di post-editing, in quanto lo sforzo cognitivo richiesto è troppo alto e risulta impossibile, dunque, dedicarsi solamente a questo impiego. Dunque, l'idea che il post-editing sia un'attività più semplice, nonché più veloce, rispetto alla traduzione è da considerarsi in molti casi scorretta, come ulteriormente dimostrato dalle risposte ottenute e precedentemente analizzate. Tale corrente di pensiero è sostenuta anche dalla partecipante 57 (post-editor, donna, 45-54 anni), secondo la quale l'errata convinzione che il post-editing richiede un minore impiego di tempo rispetto alla traduzione (da zero) porta a un grossolano sotto

pagamento, in quanto, frequentemente, il MTPE richiede lo stesso tempo, se non di più, della traduzione. La partecipante 117 (post-editor, donna, 45-54 anni) conferma che il post-editing è un lavoro faticoso a livello cognitivo, a cui può personalmente dedicare solo quattro ore al giorno anziché otto come nel caso della traduzione. Aggiunge, inoltre, che, se ci si occupa a lungo di post-editing, le soluzioni che si propongono diventano meno creative, in quanto, per l'appunto, lo sforzo cognitivo richiesto è troppo grande e l'unica volontà è quella di essere efficienti in termini di costi/velocità: il traduttore, dunque, nello svolgimento del lavoro di post-editing, si sente spesso sotto pressione.

Il post-editing, proprio per il fatto che, secondo diversi post-editor, non può essere svolto per otto ore al giorno consecutive, viene visto da alcuni traduttori come una fonte di reddito aggiuntiva, un lavoro svolto accanto al proprio impiego principale ovvero quello di traduzione (partecipante 79, donna, 55-64 anni). Altri, al contrario, ritengono il PE oggi giorno fondamentale: con una sempre maggiore presa di posizione dell'intelligenza artificiale, la sua presenza sarà essenziale, contribuendo a un cambiamento dei paradigmi della traduzione e della revisione. È quello che affermano anche la partecipante 101 (post-editor, donna, over 65) e il partecipante 76 (post-editor, uomo, over 65), il quale ammette che l'intelligenza artificiale diventerà parte integrante della nostra vita e delle nostre attività quotidiane. Il partecipante 37, infine (post-editor, uomo, over 65) afferma di utilizzare DeepL in tutte le sue traduzioni, reputando il post-editing "indispensabile" e aggiungendo che, nel suo lavoro, un mancato utilizzo del post-editing sarebbe da considerarsi un'azione irresponsabile.

La maggioranza dei partecipanti si è dichiarata insoddisfatta del metodo di remunerazione ricevuto per il post-editing, oltre al pagamento effettivo per tale impiego. Tuttavia, vi è chi ha delle idee totalmente opposte, come la partecipante 56 (post-editor, donna, 25-34 anni), che, al contrario, afferma che il post-editing, pur non essendo la mansione da lei svolta più volentieri, viene spesso ben retribuito. È semplicemente necessario, secondo la traduttrice in questione, effettuare una distinzione tra i diversi tipi di clienti, scegliendo solamente coloro che risultano maggiormente adatti ai propri interessi economici, in quanto molti clienti non hanno una corretta percezione del valore del lavoro del post-editing stesso:

I have seen jobs advertised for incredibly low prices, or had experiences where rates were incredibly low, but human level translation was expected. However, I have selected clients or continue to work with clients, who either provide reasonable compensation or understand the quality that is expected if they specify 'light' PEMT. In that sense, the PEMT work I have done and continue to do has often proven to be some of the most lucrative for me as I have separated from clients who do not fulfil this criteria<sup>17</sup>. (Partecipante 56, post-editor, donna, 25-34)

Gli insoddisfatti, ciononostante, sono numerosi. La partecipante 18 (post-editor, donna, 45-54 anni) nel suo intervento esprime quello che, almeno secondo il suo punto di vista, rappresenterebbe il problema principale del post-editing: viene affermato che i clienti del mercato della traduzione non possiedono una visione chiara delle capacità della traduzione automatica e del post-editing, e hanno come obiettivo principale e unico la riduzione dei costi. La partecipante sostiene che non esistano dei criteri precisi né per la traduzione automatica né per il post-editing riguardo al metodo di pagamento, alla sua valutazione, ecc.). Secondo la traduttrice in questione il pagamento a parola sarebbe una forma di remunerazione ridicola, in quanto non è possibile sapere quante siano le modifiche da dover apportare, sebbene siano numerose le agenzie di traduzione che utilizzano questa tipologia di retribuzione. Oltretutto, data la complessità del post-editing, la partecipante sottolinea che è fondamentale avere una corretta formazione nell'ambito e un adeguato allenamento, al fine di ridurre il più possibile i tempi, evitando di riscrivere intere frasi o sezioni.

La partecipante 23 (post-editor, donna, 55-64 anni) ha una visione altamente negativa del post-editing e, rifiutandosi di occuparsi di tale impiego, la sua esperienza nell'ambito è limitata. Esprime una grande critica al post-editing: “In my opinion the quality of the resulting translation is poorer unless the post-editor spends a lot of time for rephrasing<sup>18</sup>.” Il PE viene dunque giudicato una perdita di tempo, presentando, oltretutto, una qualità eccessivamente scadente. La partecipante 115 (post-editor, donna, 45-54 anni) dichiara che il vero problema principale del post-editing è la qualità dell'output della MT. Inoltre, secondo la sua opinione, il PE comporta meno apprendimento rispetto alla traduzione, trasformando il traduttore in una macchina: i

---

<sup>17</sup> “Ho visto lavori pubblicizzati a prezzi incredibilmente bassi o ho avuto esperienze in cui le tariffe erano incredibilmente basse, ma aspettandosi una traduzione di livello umano. Tuttavia, ho scelto clienti o continuo a lavorare con clienti che forniscono un compenso ragionevole o che comprendono la qualità che ci si aspetta se si specifica un *light post-editing*. In questo senso, il lavoro PEMT che ho svolto e che continuo a svolgere si è spesso rivelato tra i più redditizi per me, poiché mi sono allontanata dai clienti che non soddisfano questi criteri” (tradotto da me).

<sup>18</sup> “A mio parere la qualità della traduzione risultante è più scadente, a meno che il post-editor non dedichi molto tempo alla riformulazione” (tradotto da me).

clienti si aspettano che il lavoro venga svolto in tempi record, non considerando lo sforzo cognitivo richiesto dal PE e la qualità dell'output iniziale, pretendendo, tuttavia, un lavoro finale di alta qualità, pari a quella umana. Spesso, molte parti dell'output dovrebbero essere completamente riscritte, ma tale operazione non viene sempre effettuata poiché non si è normalmente pagati per farlo: la traduzione, dunque, ne risente e, conseguentemente, il prodotto finale non sarà perfetto. Pertanto, è possibile trarre profitto dal post-editing quando la qualità dell'output prodotto dal sistema di traduzione automatica è buona in partenza (partecipante 30, post-editor, donna, 55-64 anni). Quando, al contrario, la qualità risulta particolarmente scarsa, il post-editing diventa svantaggioso: è proprio per questa ragione che il pagamento su base oraria è giudicato spesso come il più adeguato.

Qualità e tempo, perciò, sono due variabili centrali nel post-editing: se l'output non è buono, il tempo impiegato e, di conseguenza, anche lo sforzo cognitivo saranno maggiori. In alcune circostanze sarebbe più vantaggioso tradurre da capo l'intero segmento o, in casi estremi, l'intero testo, come riportato pure dalla partecipante 31 (post-editor, donna, 55-64 anni). La traduttrice ribadisce che l'unico metodo di pagamento realmente equo è quello su base oraria, dal momento che il traduttore non sa come sia la qualità dell'output della MT. Secondo l'opinione della partecipante, inoltre, l'intelligenza artificiale è sopravvalutata: promette al pubblico e ai suoi utilizzatori molto più di quello che effettivamente è in grado di produrre. La qualità dell'output dipende non solo dalla tipologia di testo, ma anche dal tipo di sistema di traduzione stesso utilizzato: come sottolineato dalla partecipante 43 (post-editor, donna, 25-34 anni), vi sono alcuni sistemi che indubbiamente sono migliori di altri e che, pertanto, sono capaci di produrre un output di una qualità migliore, ottenendo, di conseguenza, un risultato finale migliore.

Anche la seguente testimonianza, riportata dalla partecipante 25 (post-editor, donna, 55-64 anni) è estremamente a sfavore del post-editing:

PE may be a source of errors because the product of machine translation is not consistent. One term may be translated in a given way in one segment but differently in another one, so the translator cannot rely on what he/she reads. [...] Besides, post-editing means having to read a greater amount of words than in normal translation, and this is very tiring for the eyes and mental concentration. It is not possible to spend full days on PE. Years ago, the post-editor was asked to just correct obvious errors, now clients expect a human-like final product out of a machine-translated text. And I find this is unfair. If a text is the product of machine translation, human intervention should be limited to

correcting obvious mistakes. It is not fair to expect the translator to make a machine product better<sup>19</sup>. (Partecipante 25, post-editor, donna, 25-34 anni)

La traduttrice parla di “ingiustizia”, di pretese troppo alte verso i post-editor, che non dovrebbero svolgere così tanto lavoro, più di quello necessario, proprio a causa della pessima qualità del prodotto della traduzione automatica. Accettare una proposta di post-editing, infatti, può essere “rischioso”, in quanto non si sa mai quanto tempo dovrà essere impiegato per questo lavoro (partecipante 47, post-editor, donna, over 65). I traduttori preferiscono occuparsi, di norma, solamente di traduzione, ma non è possibile, al giorno d’oggi, ignorare il post-editing (partecipante 46, post-editor, donna, 45-54 anni). Il fattore economico, inoltre, è una componente essenziale: è constatato che in molti casi il post-editing contribuisce a un aumento della produttività, con una riduzione dei costi. Tuttavia, il post-editing viene utilizzato dai clienti, secondo il partecipante 84 (post-editor, uomo, 35-44 anni), per pagare meno i traduttori. Il partecipante, infatti, volge un’accusa ai clienti in questione di voler intenzionalmente sottopagare il lavoro di post-editing: spesso questo lavoro viene giudicato meno importante rispetto alla traduzione e i clienti si lamentano del prezzo.

È possibile affermare, quindi, che le variabili di cui tenere conto quando si parla del pagamento del post-editing sono numerose. I clienti, tuttavia, sottovalutano tale lavoro e non ne tengono conto adeguatamente. Ne risulta che il lavoro viene sottopagato, utilizzando metodi spesso inadatti e portando a un’insoddisfazione dei traduttori. È necessaria una maggiore educazione sul tema nei confronti dei clienti stessi, al fine di ottenere equità e, pertanto, una retribuzione adeguata allo sforzo richiesto dal post-editing.

---

<sup>19</sup> Il PE può essere fonte di errori perché il prodotto della traduzione automatica non è coerente. Un termine può essere tradotto in un determinato modo in un segmento ma in modo diverso in un altro, quindi il traduttore non può fare affidamento su ciò che legge. [...] Inoltre, il post-editing implica la lettura di una quantità di parole maggiore rispetto alla traduzione normale, il che è molto faticoso per gli occhi e per la concentrazione mentale. Non è possibile dedicare intere giornate al PE. Anni fa, al post-editor si chiedeva solo di correggere gli errori più evidenti, mentre ora i clienti si aspettano un prodotto finale simile a quello umano da un testo tradotto da una macchina. E trovo che questo sia ingiusto. Se un testo fosse il prodotto di una traduzione automatica, l’intervento umano dovrebbe limitarsi a correggere gli errori più evidenti. Non è giusto aspettarsi che il traduttore migliori il prodotto di una macchina.

## Conclusione

A conclusione del lavoro di ricerca è stato possibile confermare i presupposti iniziali. Vi erano, in particolare, due premesse: si prevedeva, infatti, una scarsa conoscenza del post-editing e della sua compensazione da parte del pubblico generale, oltre a un'insoddisfazione generale dei post-editor nei confronti di propri guadagni, con la conseguente volontà di ottenere una modalità di pagamento maggiormente equa. Inoltre, vi era una terza aspettativa, ovvero che la maggioranza dei post-editor partecipanti dichiarasse di essere normalmente pagata a parola.

Come visto precedentemente, il questionario si compone principalmente di due parti. Nella prima sezione è stato richiesto a tutti i 132 partecipanti di svolgere un compito di associazione libera, ottenendo, come risultato, una lunga serie di termini, fondamentali per la delineazione della percezione del post-editing; inoltre, sono state proposte delle domande chiuse e aperte. Nella seconda sezione, invece, è stato analizzato il punto di vista dei post-editor sulla compensazione del proprio lavoro, attraverso una serie di domande chiuse, a risposta singola o multipla. Infine, è stata inserita una sezione dedicata alle riflessioni personali aggiuntive.

Di fondamentale importanza si è dimostrato il compito di associazione libera, che ha permesso di individuare il punto di vista dei partecipanti sul tema, evidenziando le differenze esistenti tra la percezione di questo lavoro e quello di traduzione. Mentre la maggioranza dei termini prodotti a partire dalla parola-stimolo *translation* sono generalmente caratterizzati da una connotazione positiva, dimostrando, inoltre, una maggiore uniformità nelle risposte registrate, quelli legati alla parola-stimolo *post-editing*, al contrario, sottolineano il divario esistente in questo mondo. Infatti, i termini relativi al post-editing possono essere raggruppati in due grandi categorie: da una parte le parole che forniscono una visione positiva di tale professione, mentre, dall'altro lato, quelle con una connotazione altamente negativa. La traduzione viene definita dai partecipanti come un processo creativo, ed è rappresentata come un ponte che permette di raggiungere e collegare lingue e culture tra loro differenti. La traduzione è descritta come un lavoro accurato che, proprio per la precisione che richiede, mostra anche una serie di difficoltà, le quali possono essere affrontate solamente se si possiedono le giuste conoscenze e abilità, grazie a un'adeguata formazione. Inoltre, è percepita come una professione interessante, utile, essenziale, e in alcuni casi divertente. "Traduzione", per

diversi partecipanti, è anche sinonimo di eccellenza, oltre che di professionalità, fiducia, passione e libertà. Oltretutto, permette di ottenere una paga dignitosa. Alcune persone definiscono tale impiego come un lavoro dispendioso in termini di tempo, mentre altre affermano che, in determinate circostanze, risulta addirittura più veloce del post-editing. Infine, alcuni partecipanti hanno espresso la propria preoccupazione, dettata dalla sensazione che la traduzione (umana), oggi, venga ignorata, rischiando di essere surclassata dalla traduzione automatica.

Per quanto riguarda i termini prodotti a partire dalla parola-stimolo *post-editing*, invece, vi è una minore uniformità nei risultati: più precisamente, sono stati maggiormente citati quei sostantivi e aggettivi che descrivono in cosa consiste il lavoro di post-editing stesso, quali *revision*, *correction*, *adjustment*, ecc. Tali parole esprimono, infatti, che si tratta di un lavoro di revisione dell'output prodotto da un sistema di traduzione automatica, effettuato individuando e correggendo gli errori presenti. Tuttavia, per comprendere l'effettiva percezione di questa professione, risultano più interessanti tutti quei termini che, sebbene presentano un numero di occorrenze inferiore, permettono di raccogliere le opinioni personali dei singoli partecipanti: i post-editor stessi, in particolare, possono fornire una visione interna del fenomeno, essendo questa la propria professione. Come accennato, le parole citate sono tra loro altamente discostanti, ed evidenziano le enormi differenze esistenti nella percezione del post-editing, soprattutto tra i post-editor stessi. Infatti, vi sono dei sostenitori di tale professione, che la definiscono moderna e veloce, in quanto permette di risparmiare sui tempi di traduzione e consente di correggere facilmente gli errori presenti. Secondo alcuni, inoltre, il post-editing è interessante, innovativo, efficiente e addirittura indispensabile. Tra gli oppositori, invece, vi sono coloro che reputano il lavoro di post-editing estremamente noioso, in quanto ripetitivo e poco creativo. Sono molti i traduttori che sostengono che la qualità dell'output prodotto dalla *machine translation* è troppo scarsa, e per tale ragione dev'essere dedicato molto tempo alla sua correzione. Il post-editing viene percepito come un'attività difficile e frustrante, piena di ostacoli e di cui non vale la pena occuparsi. Alcuni partecipanti affermano che si tratta di un lavoro superficiale e inadatto, in quanto non può essere considerato come una vera tipologia di traduzione, e che, come sottolineato da un traduttore, fa sentire i post-editor come degli "schiavi" (il partecipante in questione ha infatti utilizzato la parola *slave* per descrivere

il post-editing). Vi sono, inoltre, due correnti di pensiero riguardo alla sua compensazione: da una parte coloro che ritengono il post-editing un mezzo economico, che permette di risparmiare efficacemente sui costi; dall'altra parte, invece, c'è chi sostiene che è fin troppo sottopagato, proprio a causa del risparmio sui costi esistente.

Successivamente, sono state esaminate le risposte riguardanti le domande aperte e chiuse. Sebbene diverse persone tra il pubblico generale abbiano affermato di aver sentito parlare del post-editing prima del questionario sottopostogli, quasi nessuno ha saputo fornire una precisa indicazione sui metodi di pagamento esistenti per tale lavoro. I partecipanti, facenti parte del pubblico generale, potrebbero rappresentare, infatti, al momento attuale o futuro, dei possibili clienti: per tale motivo una scarsa concezione del post-editing e della sua remunerazione non sono considerati dei fattori positivi. Come segnalato dall'88% dei post-editor partecipanti allo studio, in effetti, il pubblico e, conseguentemente, i clienti non hanno generalmente una corretta percezione di questo lavoro e della sua compensazione. Tale disinformazione provoca, molte volte, il ricorso a un metodo di pagamento poco consono e inadatto, portando a condizioni di lavoro vantaggiose solamente per i clienti stessi, il cui obiettivo principale è la riduzione dei costi, attraverso la diminuzione dei tempi normalmente impiegati per la traduzione, con la richiesta di ottenere, tuttavia, un risultato di una qualità comparabile a quella umana.

Dall'analisi delle risposte fornite dai post-editor sul proprio grado di soddisfazione rispetto al lavoro svolto e al guadagno ottenuto, risulta che sono numerosi coloro che si sono dichiarati insoddisfatti. In quasi tutte le risposte, infatti, circa la metà degli intervistati ha dimostrato di non essere soddisfatta del sistema di remunerazione ricevuto. Come previsto, la maggioranza dei post-editor viene pagata a parola, ma molti preferirebbero un sistema di compensazione più equo: infatti, di frequente il pagamento a parola non tiene conto della complessità del lavoro di post-editing, che non consiste nella semplice correzione degli errori presenti nell'output, ma in una lunga serie di fasi, quali la lettura dell'output stesso e del testo sorgente, l'individuazione degli errori e la loro correzione, la rilettura del lavoro finale e la riconsegna al cliente. Inoltre, l'output prodotto dai sistemi di traduzione neurali, sebbene sia nettamente migliorato nel corso degli ultimi anni, deve essere ulteriormente perfezionato. Infatti, non è ancora possibile affermare il raggiungimento di una qualità pari o molto simile a quella umana, in modo particolare con determinate coppie linguistiche. La scarsa qualità dell'output

rappresenta uno dei principali problemi del post-editing e contribuisce a rendere il suo processo nettamente più lungo e complesso, richiedendo, oltretutto, un maggiore sforzo cognitivo. In alcune circostanze, i post-editor preferiscono ritradurre un'intera porzione di testo anziché effettuare il post-editing, poiché la qualità risulta estremamente bassa e la revisione richiederebbe troppo tempo, influenzando pesantemente sulla propria produttività e, di conseguenza, sui guadagni ottenuti.

Il tempo e la qualità dell'output della MT, dunque, sono due variabili fondamentali; tuttavia, la retribuzione a parola normalmente non ne tiene sufficientemente conto. Infatti, come sottolineato da alcuni partecipanti, sarebbe preferibile un pagamento su base oraria, che permette ai post-editor di lavorare con meno pressioni e di tenere realmente conto dei diversi elementi che influiscono sul loro lavoro. Non raramente i traduttori, occupandosi di post-editing, affermano di sentirsi sotto pressione, in quanto devono svolgere il lavoro nel minor tempo possibile al fine di non avere delle ripercussioni sui propri guadagni, assicurando, ciononostante, una qualità finale pari a quella umana. Inoltre, è necessario sottolineare che il MTPE non è adatto a tutti i tipi di testo. Tuttavia, i clienti non tengono conto di questo importante fattore, ottenendo un output MT di pessima qualità, in quanto il contenuto del testo sorgente non è adatto a essere tradotto con un sistema di traduzione automatica. Ciononostante, al giorno d'oggi all'interno del mercato della traduzione vi è una sempre maggiore richiesta di post-editing: infatti, analizzando le motivazioni fornite dai post-editor riguardo alla scelta di praticare tale lavoro, la maggioranza dei partecipanti ha dichiarato che si tratta di una scelta inevitabile, essendo uno dei lavori più richiesti dal mercato target. Alcuni post-editor hanno sottolineato il fatto che si tratta, in molti casi, di una scelta obbligata, al fine di ottenere il lavoro.

Procedendo poi, con l'analisi delle risposte relative al metodo di pagamento ritenuto migliore dai partecipanti, è possibile affermare con certezza che il 44.1% dei partecipanti preferirebbe essere pagato con un pagamento su base oraria, giudicato come il metodo di pagamento più equo per il post-editing. Tuttavia, una considerevole porzione di post-editor continua a preferire il pagamento a parola (33.8%). In effetti, il calcolo del pagamento a parola non viene sempre effettuato nello stesso modo, in quanto alcuni clienti tengono conto anche della distanza di modifica o, addirittura, effettuano un pagamento sia per le parole da modificare che per quelle del testo

sorgente, come accade nell'azienda Translated (precedentemente citato nel corso del terzo capitolo). In molte altre circostanze, però, il pagamento ricevuto non soddisfa le aspettative dei post-editor. Infatti, molti affermano che solamente il pagamento su base oraria potrebbe realmente tenere conto di tutte le caratteristiche del testo, oltre che dell'effettivo sforzo cognitivo richiesto da ognuno dei progetti di post-editing affidati.

Il problema alla base di questa discussione, che si protrae ormai da diversi anni, è la scarsa e scorretta informazione dei clienti. Innanzitutto, dovrebbero essere fornite delle linee guida più specifiche riguardanti i metodi di pagamento accettati: il post-editing, infatti, è ormai un'attività esistente da diversi anni e se ne sono comprese le difficoltà, dunque i punti di debolezza, oltre a quelli di forza. Infatti, se utilizzato nella maniera corretta e con gli strumenti giusti, il post-editing consente un reale aumento della produttività e dimezzamento dei tempi. È necessario, però, garantire una maggiore tutela ai traduttori che svolgono questo impiego, data la sua estrema richiesta da parte del mercato della traduzione stesso. Così facendo, si potrebbe arrivare a stabilire un metodo di pagamento unico ed equo, valido per tutti i post-editor e tutti i clienti. Al contrario, al momento attuale, vi è una situazione di disinformazione generale dei clienti che, nel maggioranza dei casi, sono proprio coloro a cui spetta la decisione del metodo di remunerazione da applicare e le condizioni secondo le quali applicarlo. Il lavoro di post-editing potrebbe, in tal modo, ottenere una considerazione maggiore ed essere più apprezzato. Inoltre, dovrebbero sussistere delle indicazioni precise riguardo alle tipologie di testo e contenuto delle quali si può effettuare il post-editing, su cui i clienti dovrebbero essere informati. I post-editor, d'altro canto, dovrebbero, per quanto possibile, rifiutarsi di accettare pagamenti troppo bassi, facendo comprendere al cliente il valore delle proprie capacità e del proprio lavoro. Se il futuro della traduzione risiede proprio nel post-editing, vi è la necessità di ottenere un pagamento dignitoso, al fine di avvicinare nuovi futuri traduttori (e post-editor) al mondo della traduzione e non, al contrario, provocare il loro allontanamento, in quanto spaventati dalle condizioni lavorative proposte.



## Appendice

### **A fair rate for machine translation post-editing**

Dear participant! First of all, thank you for participating in the following survey. This project is part of my final dissertation at the University of Padua (Italy), where I am currently enrolled in the MA course of Modern Languages for International Communication and Cooperation. This research aims to study how people perceive machine translation post-editing work and especially what is the fairest pricing method for it, focusing, then, on the direct point of view of post-editors. For this purpose, we have prepared an anonymous questionnaire that we invite you to fill out.

Those who participate in the research can withdraw at any time without giving any justification, obtaining non-use of the data.

### **Data processing information form**

We would like to remind you that the survey is anonymous and that the data will only be used for scientific research purposes. The data will not investigate the characteristics of the individual participant but will be processed with statistical techniques exclusively at the group level. However, please answer all questions individually in the order in which they are presented. All materials will be processed and stored in full respect of privacy and anonymity according to the Italian Legislative Decree No. 196 of 30.06.2003 and subsequent amendments and int. "Codice in materia di protezione dei dati personali" (the so-called Privacy Code).

Should they request it, participants have the right to access their data.

### **Consent to participate in research and data processing**

By filling out this questionnaire, you agree with the data protection regulations in the data privacy statement.

### **Contacts**

Greta Maniero

[greta.maniero.2@studenti.unipd.it](mailto:greta.maniero.2@studenti.unipd.it)

### **Section 1**

1. Please write down what words might come to mind if you were asked to think of the word "translation". (*write between 2 and 5 nouns or adjectives*)

2. Please write down what words might come to mind if you were asked to think of the word "post-editing". (*write between 2 and 5 nouns or adjectives*)

---

3. Before this questionnaire, have you ever heard of post-editing?

*Mark only one oval.*

- Yes
- No
- I am not sure about it

4. If you have heard about post-editing before, where have you heard about it?

*Check all that apply.*

- At work
- At school/University
- On the Internet
- On social media
- In some book
- On television
- In some newspapers/magazine
- In some movie
- Talking with my family/friends
- Can't remember where
- Other (Please fill in)

5. Do you think the payment methods are the same for post-editing and translation?

*Mark only one oval.*

- Yes
- No
- I am not sure

6. Are You aware of the existing payment methods for post-editing?

*Mark only one oval.*

- Yes
- No
- I am not sure

7. Can you please mention the existing payment methods for post-editing you are aware of (or that you think you are aware of)?
- 

8. What is your professional occupation?

*Mark only one oval.*

- Unemployed
- Student
- Retired
- Entrepreneur
- Self-employed
- Teacher/Professor
- Office worker
- Technician
- Serviceperson
- Worker
- Other (Please fill in)

9. Do you work, or have you ever worked, as a post-editor? *(if is yes, you can answer the questions in the next section; otherwise, you can skip it and go to the last section)*

*Mark only one oval.*

- Yes
- No

## **Section 2: Post-editing and fair pricing**

The following section is for professional translators/post-editors only. Please respond **only** if you belong to this category.

10. What are the reasons why you decided to post-edit?

*Check all that apply.*

- Demand of the target market
- Productivity increase
- Increase in profits
- Cost reduction
- Time saving

- Fair payment method
- Other (Please fill in)

11. Do you think the general public has a correct perception of post-editing and its compensation in terms of money?

*Mark only one oval.*

- Yes
- No
- I don't know

12. How do you usually get paid when you do post-editing?

*Mark only one oval.*

- Payment per word
- Payment per hour
- Edit-distance
- Payment per project
- Other (Please fill in)

13. Are you satisfied with the payment method by which you get paid?

*Mark only one oval.*

- Yes
- No

14. According to your opinion, what is the most suitable payment model when post-editing?

*Mark only one oval.*

- Payment per word
- Payment per hour
- Edit-distance
- Payment per project
- Other (Please fill in)

15. Are rates adequate for the effort required for PE? Think about what you can earn in a day only doing PE.

*Mark only one oval.*

- Yes
- No

- I don't know

16. In your opinion, how fair are post-editing rates compared to the time spent?

*Mark only one oval.*

- After a full day only post-editing, I make more money than if I was translating.
- After a full day only post-editing, I make less money than if I was translating.
- After a full day only post-editing, I make the same money as if I was translating.

17. Please indicate your level of agreement with the following statements:

*Mark only one oval per row.*

- a. The per word rate is the fairest method of payment.
  - Strongly agree
  - Somewhat agree
  - Somewhat disagree
  - Strongly disagree
  - I don't know
- b. The per hour rate is the fairest method of payment for MTPE.
  - Strongly agree
  - Somewhat agree
  - Somewhat disagree
  - Strongly disagree
  - I don't know
- c. I am satisfied with my earnings from post-editing.
  - Strongly agree
  - Somewhat agree
  - Somewhat disagree
  - Strongly disagree
  - I don't know
- d. Post-editing reduces my costs.
  - Strongly agree
  - Somewhat agree

- Somewhat disagree
  - Strongly disagree
  - I don't know
- e. Post-editing is a waste of time and money.
  - Strongly agree
  - Somewhat agree
  - Somewhat disagree
  - Strongly disagree
  - I don't know
- f. The ratio of time spent doing post-editing to income received is adequate.
  - Strongly agree
  - Somewhat agree
  - Somewhat disagree
  - Strongly disagree
  - I don't know

**Section 3: Personal information**

18. What gender do you identify with?

*Mark only one oval.*

- Woman
- Man
- Non-binary
- Prefer not to answer
- Other (Please fill in)

19. Age

*Mark only one oval.*

- Under 18
- 18-24
- 25-34
- 35-44
- 45-54
- 55-64
- 65+

20. What is your highest level of education?

*Mark only one oval.*

- Middle school diploma
- Professional school diploma
- High school diploma
- Bachelor's degree
- Master's degree
- PhD/Specialization course
- Other (Please fill in)

21. If you want to add something more to the topic:

---



## Bibliografia

- ALPAC, *Languages and Machines: Computers in Translation and Linguistics. A report by the Automatic Language Processing Advisory Committee, Division of Behavioral Sciences*, Washington, D.C., National Academy of Sciences, National Research Council, 1966
- Arnhold, P., Grizzo, S., Hansen-Schirra, S., Nitzke, J., Porsiel, J., Schaeffer, M., Schneider, M., *Maschinelle Übersetzung und Post-Editing*, in “Best Practices - Übersetzen und Dolmetschen”, Ottman, A. (a cura di), Berlin, BDÜ-Fachverlag, pp. 221–249, 2017
- Asmuth, V., *Edit-Distance Based Compensation for Machine Translation*, in “Maschinelle Übersetzung für Übersetzungsprofis” Porsiel, J. (a cura di), pp. 368-374, 2020
- Bowker, L., *Terminology*, in “Routledge Encyclopedia of Translation Studies. Second Edition”, Baker, M., Saldanha, G. (a cura di), London and New York, Routledge. Taylor and Francis Group, pp. 286-290, 2009
- Bowker, L., *Fit-for-purpose translation*, in “The Routledge Handbook of Translation and Technology”, O’Hagan, M. (a cura di), Abingdon, Oxon, Routledge, pp. 453-468, 2020
- Bowker, L., Fisher, D., *Computer-aided translation*, in “Handbook of Translation Studies”, Vol. 1, Gambier, Y., Van Doorslaer, L. (a cura di), Amsterdam, John Benjamins Publishing Company, p. 60-65, 2010
- Brondi, S., Neresini, F., *Studying the emergence of a new social representation: Changes in thinking about nanotechnologies in early 21st-century Italy*, in “European Journal of Social Psychology”, Vol. 48, No. 6, pp. 815-833, 2018
- Burchardt, A., Macketanz, V., Dehdari, J., Heigold, G., Peter, J., Williams, P., *A Linguistic Evaluation of Rule-Based, Phrase-Based, and Neural MT Engines*, in “Prague Bulletin of Mathematical Linguistics”, No. 108, pp. 159-170, 2017
- Cadwell, P., O’Brien, S., Teixeira, C., *Resistance and accommodation: Factors for the (non-) adoption of machine translation among professional translators*, in “Perspectives”, Vol. 26, No. 3, 301–321, 2018
- Canfora, C., Ottmann, A., *Risiken und Haftungsfragen bei neuronaler maschineller Übersetzung*, in “Maschinelle Übersetzung für Übersetzungsprofis”, Porsiel, J. (a cura di), Berlino, BDÜ-Fachverlag, pp. 171-184, 2020a
- Canfora, C., Ottmann, A., *Translation risk management – error prevention and near-misses management*, in “Translation Quality in the Age of Digital Transformation”, Dalla-Zuanna, J., Kurz, C. (a cura di), Berlino, BDÜ-Fachverlag, pp. 136-153, 2020b
- Castilho, S., Moorkens, J., Gaspari, F., Calixto, I., Tinsley, J., Way, A., *Is Neural Machine Translation the New State of the Art?*, in “The Prague Bulletin of Mathematical Linguistics”, N. 108, pp. 109-120, 2017
- Dany, L., Urdapilleta, I., Lo Monaco, G., *Free associations and social representation: some reflections on rank-frequency and importance-frequency methods*, in “Quality & Quantity: International Journal of Methodology”, Vol. 49, No. 2, Springer, pp. 439-507, 2014

- Do Carmo, F., Moorkens, J., *Differentiating editing, post-editing and revision*, in “Translation Revision and Post-Editing: Industry Practices and Cognitive Processes”, Koponen, M., *et al.* (a cura di), Londra, Routledge, pp. 35-49, 2021
- EMT, Expert Group, *Competences for professional translators, experts in multilingual and multimedia communication*, Bruxelles, 2009
- Forcada, M. L., *Machine translation today*, in “Handbook of Translation Studies”, Vol. 1, Gambier, Y., Van Doorslaer, L. (a cura di), Amsterdam, John Benjamins Publishing Company, p. 215-223, 2010
- Ginovart Cid, C., Colominas Ventura, C., *The MT post-editing skill set: Course descriptions and educator’s thoughts*, in “Translation Revision and Post-editing: Industry Practices and Cognitive Processes”, Koponen, M., *et al.* (a cura di), Londra, Routledge, pp. 226-246, 2021
- Grizzo, S., *Bezahlmodelle für Post Editing*, in “Maschinelle Übersetzung für Übersetzungsprofis” Porsiel, J. (a cura di), pp. 363-367, 2020
- Guerberof Arenas, A., *Productivity and quality in MT post-editing*, in “Beyond Translation Memories: New Tools for Translators Workshop”, MT Summit, Ottawa, Canada, 2009
- Guerberof Arenas, A., *What do professional translators think about post-editing?*, in “The Journal of Specialised Translation”, Vol. 19, pp. 75-95, 2013
- Guerberof Arenas, A., *Correlations between Productivity and Quality When Post-Editing in a Professional Context*, in “Machine Translation”, Vol. 28, No. 3-4, pp. 165–186, 2014
- Guerberof Arenas, A., Moorkens, J., *Machine translation and post-editing training as part of a master’s Programme*, The Journal of Specialised Translation, Vol. 31, pp. 217-238, 2019
- Han, B., *Translation, from Pen-and-Paper to Computer-Assisted Tools (CAT Tools) and Machine Translation (MT)*, in “Proceedings”, Vol. 63, No. 56, 2020
- Hearne, M., Way, A., *Statistical Machine Translation: A Guide for Linguists and Translators*, in “Language and Linguistic Compass”, Vol. 5, no. 5, pp. 205-226, 2011
- Hutchins, W. J., *The first decades of machine translation: Overview, Chronology, Sources*, in “Early Years in Machine Translation”, Hutchins, W. J. (a cura di), Amsterdam, John Benjamins Publishing Company, pp. 1-16, 2000
- Hutchins, W. J., *Machine Translation: A Brief History*, in “Concise History of the Language Sciences. From the Sumerians to the Cognitivists”, Koerner, E. F., Asher, R. E. (a cura di), Oxford, Pergamon, pp. 431-451, 1995
- International Organization for Standardization, 18587, DIN ISO, Translation services — post-editing of machine translation output — requirements, 2018
- International Organization for Standardization, 31000, ISO, International Standard: Risk management: Principles and guidelines, 2009
- Kamocki, P., O’Regan, J., Stauch, M., *All your data are Be-long to us. European perspectives on privacy issues in ‘free’ online Machine translation services*, in “IFIP

- international summer school on privacy and identity management”, pp. 265–280, 2015
- Kamocki, P., Stauch, M., “*Cover this data that I cannot see*”: *Privacy by Design in Machine Translation*, in “Maschinelle Übersetzung für Übersetzungsprofis”, Porsiel, J. (a cura di), Berlino, BDÜ Fachverlag, pp. 42-57, 2020
- Kanavos, P., Kartsaklis, D., *Integrating Machine Translation with translation memory: A practical approach*, in “Proceedings of the Second Joint EM+/CNGL Workshop Bringing MT to the User: Research on Integrating MT in the Translation Industry”, pp. 11-20, 2010
- Kenny, D., *Human and machine translation*, in “Machine Translation for Everyone. Empowering users in the age of artificial intelligence”, Kenny, D. (a cura di), Berlino, Language Science Press, pp. 23-50, 2022
- King, M., Perschke, S.; *EUROTRA and its objectives*, in “Multilingua”, Vol.1, no. 1, pp. 27-32, 1982
- Koehn, P., Knowles, R., *Six Challenges for Neural Machine Translation*, in “Proceedings of the 1<sup>st</sup> Workshop on neural machine translation”, Vancouver, BC, Canada, pp. 28-39, 2017
- Koehn, P., Knowles, R., Sanchez Torron, M., *Neural Interactive Translation Prediction*, in “Maschinelle Übersetzung für Übersetzungsprofis. Sammelband”, Porsiel, J. (a cura di), BDÜ-Fachverlag, Berlino, pp. 116- 144, 2020
- Kontinen, K., Salmi, L., Koponen, M., *Revision and post-editing competences in translator education*, in “Translation Revision and Post-editing: Industry Practices and Cognitive Processes” Koponen, M., et al. (a cura di), Londra, Routledge, pp. 187-202, 2021
- Koponen, M., Salmi, L., *On the correctness of machine translation: A machine translation post-editing task*, in “The Journal of Specialised Translation”, Issue 23, pp. 118-136, 2015
- Koponen, M., *Is machine translation post-editing worth the effort? A survey of research into post-editing and effort*, in “The Journal of Specialised Translation”, Issue 25, 2016a
- Koponen, M., *Machine Translation Post-editing and Effort: Empirical Studies on the Post-editing Process*, Helsinki, Unigrafia, 2016b
- Koponen, M., Mossop, B., Robert, I., Scocchera, G., *Translation Revision and Post-editing: Industry Practices and Cognitive Processes*, Londra, Routledge, 2021
- Läubli, S., Orrego-Carmona, D., *When Google Translate is better than Some Human Colleagues, those People are no longer Colleagues*, in “Translating and the Computer”, London, pp. 59-69, 2017
- Macketanz, V., Avramidis, E., Burchardt, A., Helcl, J., Srivastava, A., *Machine translation: Phrase-based, rule-based and neural approaches with linguistic evaluation*, in “Cybernetics and Information Technologies”, Vol. 17, No 2, pp. 28–43, 2017
- Massardo, I., Van der Meer, J., O’Brien, S., Hollowood, F., Aranberri, N., Drescher, K., *MT Post-editing Guidelines*, Amsterdam, TAUS Signature Editions, 2016

- Melby, A. K., *Future of machine translation. Musings on Weaver's memo*, in "The Routledge Handbook of Translation and Technology", O'Hagan, M. (a cura di), New York, Routledge, pp. 419-436, 2020
- Mitchell, L., Roturier, J., O'Brien, S., *Community-based post-editing of machine-translated content: Monolingual vs. Bilingual*, in "Proceedings of the MT Summit Conference", pp. 35-43, 2013
- Moorkens, J., O'Brien, S., da Silva, I.A.L. *et al.*, *Correlations of perceived post-editing effort with measurements of actual effort*, in "Machine Translation", Vol. 29, pp. 267-284, 2015
- Moorkens, J., O'Brien, S., *Assessing user interface needs of post-editors of machine translation*, in "Human Issues in Translation Technology", Kenny, D. (a cura di), Vol. 1, New York, Routledge, pp. 109-130, 2017
- Moorkens, J., *Ethics and machine translation*, in "Machine translation for everyone: Empowering users in the age of artificial intelligence", Kenny, D. (a cura di), pp. 121-140, 2022
- Nitzke, J., Hansen-Schirra, S., Canfora, C., *Risk management and post-editing competence*, in "The Journal of Specialised Translation", Issue 31, pp. 239-259, 2019.
- Nitzke, J., Hansen-Schirra, S., *A short guide to post-editing*, Berlin, Language Science Press, 2021
- Nitzke, J., Gros, A., *Preferential Changes in Revision and Post-editing*, in "Translation Revision and Post-editing. Industry Practices and Cognitive Processes", Koponen, M., Mossop, B., Robert, I., Scocchera, G. (a cura di), pp. 21-34, 2021
- O'Brien, S., *Teaching Post-editing: A Proposal for Course Content*, in "Proceeding of the 6th EAMT Workshop: Teaching Machine Translation", Manchester, Regno Unito, pp. 99-106, 2002
- O'Brien, S., *Towards predicting post-editing productivity*, in "Machine Translation", Vol. 25, pp. 197-215, 2011
- O'Brien, S., *Post-editing*, in "Handbook of Translation Studies", Gambier, Y., *et al.* (a cura di), Vol. 5, Amsterdam, John Benjamins, pp. 177-183, 2021
- O'Hogan, M., *Computer-aided-translation (CAT)*, in "Routledge Encyclopedia of Translation Studies. Second Edition", Baker, M., Saldanha, G. (a cura di), London and New York, Routledge. Taylor and Francis Group, pp. 48-51, 2009
- Ortega, J. E., Turchi, M., Sánchez-Martínez, F., Negri, M., *Improving Translations by Combining Fuzzy-Match Repair with Automatic Post-Editing*, in "Proceedings of MT Summit XVII", Vol. 1, Dublin, pp. 256-266, 2019
- Pérez-Ortiz, J. A., Forcada, M. L., Sánchez-Martínez, F., *How neural machine translation works*, in "Machine Translation for Everyone. Empowering users in the age of artificial intelligence", Kenny, D. (a cura di), Berlino, Language Science Press, 2022 pp. 141-164, 2022
- Poibeau, T., *Machine Translation*, Cambridge, Massachusetts, The MIT Press, 2017
- Rico, C., Torrejón, E., *Skills and Profile of the New Role of the Translator as MT Post-editor*, in "Revista Tradumàtica", No 10, pp. 166-178, 2012

- Sakamoto, A., *Why do many translators resist post-editing? A sociological analysis using Bourdieu's concepts*, in "The Journal of Specialised Translation", Vol. 31, pp. 201-216, 2019
- Schwartz, L., *The history and promise of machine translation*, in "Innovation and Expansion in Translation Process Research", Lacruz, I., Jääskeläinen, R. (a cura di), Amsterdam, John Benjamins Publishing Company, pp. 161-190, 2018
- Stasimioti, M., Sosoni, V., *Undergraduate Translation Students' Performance and Attitude vis-à-vis Machine Translation and Post-editing: Does Training Play a Role?*, in "Proceedings of the Translating and the Computer Conference 41", pp. 125-136, 2019
- Toma, P., *From Serna to Systran*, in "Early Years in Machine Translation", Hutchins, W. J. (a cura di), Amsterdam, John Benjamins Publishing Company, pp. 135-146, 2000
- Van Genabith, J., *Neural Machine Translation*, in "Maschinelle Übersetzung für Übersetzungsprofis. Sammelband", Porsiel, J. (a cura di), Berlino, BDÜ-Fachverlag, pp. 59-115, 2020
- Vardaro, J., Schaeffer, M., *Translation Quality and Error Recognition in Professional Neural Machine Translation Post-Editing*, in "Informatics", Vol. 6, No. 41, 2019

### Sitografia

- Cambridge Dictionary Online: <https://dictionary.cambridge.org/>
- DeepL.com: <https://www.deepl.com/privacy>
- Schelfhout, C., *Edit-distance: not a miracle cure*, 2022 <https://www.rws.com/blog/edit-distance/>
- Scribbr.de: <https://www.scribbr.de/hausarbeit/gender-hinweis-hausarbeit-vorlage/>
- Slator.com: <https://slator.com/translated-marco-trombetti-time-to-edit-as-proxy-for-ai-singularity/>
- TAUS.com: <https://www.taus.net/company/about-us>
- Translated.com: <https://translated.com/speed-to-singularity>



## Zusammenfassung

Der Inhalt dieser Abschlussarbeit handelt von der Untersuchung und Analyse der Vergütung von Post-Editing. Der Begriff „Post-Editing“ bezieht sich auf die Überprüfung des von einem maschinellen Übersetzungssystem produzierten Outputs durch einen menschlichen Übersetzer<sup>20</sup>, der als "Post-Editor" bezeichnet wird und in diesem Bereich ausgebildet ist. Da das Post-Editing eine andere Tätigkeit als die Übersetzung ist, gibt es auch unterschiedliche Bezahlmodelle. Diese Frage ist derzeit umstritten, da sich noch keine einheitliche und allgemeingültige Vergütungsmethode durchgesetzt hat. Das Hauptziel dieser Abschlussarbeit besteht also darin, die Wahrnehmung von Post-Editing sowohl bei den Post-Editoren selbst als auch bei der allgemeinen Öffentlichkeit zu untersuchen. Das allgemeine Publikum besteht aus Studenten, Unternehmern, Arbeitnehmern, Angestellten usw. Diese können potenzielle Kunden sein oder mit Post-Editing oder maschineller Übersetzung in Berührung kommen. Das Ziel ist daher, ihre Meinung zu diesem Thema zu verstehen und möglicherweise die beste Zahlungsmethode für Post-Editing zu bestimmen, d. h. ein gerechtes und angemessenes Bezahlmodell, das die verschiedenen Variablen von Post-Editing berücksichtigt, wie z. B. die Qualität der vom maschinellen Übersetzungssystem erzeugten Ausgabe, das Sprachenpaar, den Inhalt des Textes und die für das Lesen des Ausgangstextes und der endgültigen Arbeit erforderliche Zeit.

Hinter der Wahl des Themas lag ein Interesse an der Welt der Übersetzung zugrunde, das meinem Universitätsstudium entsprach. Außerdem war ich neugierig, mehr über ein so kontroverses Thema zu erfahren, da es noch keine gültige Lösung dafür gibt. Ich persönlich fand es sehr interessant, die Antworten der Post-Editoren und der breiten Öffentlichkeit zu diesem Thema zu analysieren. Von besonderer Bedeutung waren die ausführlichen Kommentare der Post-Editoren im letzten Abschnitt des Fragebogens, die wichtige Denkanstöße lieferten. Die in dieser Abschlussarbeit vorgeschlagene Analyse stellt, wenn auch in kleinem Rahmen, das allgemeine Denken der Post-Editoren zum Thema Vergütung dar und ermöglicht ein tieferes Verständnis der damit verbundenen Probleme. Darüber hinaus besteht die Hoffnung, dass die von Post-Editoren geleistete

---

<sup>20</sup> **Gender-Hinweis:** Zur besseren Lesbarkeit wird in dieser Abschlussarbeit das generische Maskulinum verwendet. Die in dieser Arbeit verwendeten Personenbezeichnungen beziehen sich – sofern nicht anders kenntlich gemacht – auf alle Geschlechter (Scribbr.com).

Arbeit und die erforderlichen kognitiven Anstrengungen endlich mehr Anerkennung finden und die Kunden besser über dieses Thema informiert werden.

Die Arbeit ist in drei Hauptkapitel unterteilt. Das Thema des ersten Kapitels ist die maschinelle Übersetzung (MÜ): Die maschinelle Übersetzung ist eine Art von Übersetzung, die durch ein Computerprogramm automatisch erstellt wird. Unter bestimmten Umständen ist in der MÜ ein menschliches Eingreifen in Form von Pre- oder Post-Editing erforderlich. Bei der maschinellen Übersetzung wird, wie bei der menschlichen Übersetzung, der Ausgangstext in eine bestimmte Zielsprache übertragen; die Produktion des Textes in der Zielsprache erfolgt jedoch automatisch unter Verwendung eines Übersetzungssystems. Sie gilt als eine der ersten nicht-numerischen Anwendungen von Computern, die nach dem Zweiten Weltkrieg entwickelt wurden. Das erste Kapitel fasst die wichtigsten Ereignisse der Geschichte der MÜ zusammen. Die ersten Experimente auf dem Gebiet der MÜ gehen auf die erste Hälfte der 1930er Jahre zurück, als die Ingenieure Georges Artsrouni und Pëtr Trojanskij erstmals Patente für die damals genannten *translating machines* anmeldeten. Am Anfang wurde die MÜ *mechanical translation* bezeichnet, weil der Übersetzungsprozess sehr kompliziert war. Jedoch die Forschung im Bereich der maschinellen Übersetzung begann offiziell 1949 dank der Ideen und Entdeckungen von Weaver, die er in seinem Memorandum veröffentlichte. In dieser Arbeit werden die wichtigsten Etappen bis hin zu der heute erreichten Entwicklung beschrieben. Während die Übersetzungssysteme am Anfang nur von internationalen Organisationen und Regierungen genutzt wurden, kann man heute sagen, dass die maschinelle Übersetzung zu einem Massenkonsumgut geworden ist: Jeder kann nämlich die Vorteile online verfügbarer kostenloser maschineller Übersetzungssysteme nutzen. Google Translate, zum Beispiel, ist heutzutage das berühmteste Übersetzungssystem der Welt. Die meisten der erwähnten historischen Ereignisse sind den Schriften von Hutchins (1995 und 2000) entnommen, in denen über die anfänglichen Probleme und Herausforderungen berichtet wird, die sich im Laufe der Entwicklung der maschinellen Übersetzung stellten.

In der mehr als 70-jährigen Geschichte der maschinellen Übersetzung folgten drei Übersetzungssysteme aufeinander: die regelbasierte maschinelle Übersetzung (RBMÜ – *rule based machine translation*), die statistische maschinelle Übersetzung (SMÜ –

*statistical machine translation*) und die neuronale maschinelle Übersetzung (NMÜ – *neural machine translation*). Die RBMÜ-Systeme waren zwischen den 1950er und 1980er Jahren die dominierenden Systeme. Die 1990er Jahre markierten eine Übergangsphase zwischen RBMÜ- und SMÜ-Systemen. Ab den frühen 2000er Jahren überholten die SMÜ-Systeme ihre Vorgänger. Seit 2018 gelten neurale Systeme als die fortschrittlichsten. In dieser Abschlussarbeit werden die wichtigsten Gemeinsamkeiten und Unterschiede zwischen den drei Systemen analysiert: Insbesondere ist das RBMÜ-System regelbasiert, während die SMÜ- und NMÜ-Systeme datengesteuerte Systeme sind und die Ergebnisse der Trainingsdatenanalyse nutzen. NMÜ ist hauptsächlich eine Technologie, die auf künstlichen neuronalen Netzen basiert und aufgrund der immer besser werdenden künstlichen Intelligenz in den letzten Jahren erhebliche Fortschritte gemacht hat.

In der Arbeit werden dann die wichtigsten Unterschiede zwischen MÜ und CAT analysiert: Mit diesem zweiten Begriff meint man eine "computergestützte Übersetzung" (CAT). Diese Art von Übersetzung entstand in den 60er Jahren mit dem Ziel, den (menschlichen) Übersetzer bei seiner Arbeit zu unterstützen. Insbesondere betonte Martin Kay in den 80er Jahren die Notwendigkeit, die Aufmerksamkeit wieder auf die Menschen zu lenken, da man sich in diesem historischen Moment nur auf die maschinelle Übersetzung konzentrierte. Heutzutage sind die CAT-Systeme zu einem unverzichtbaren Werkzeug in der täglichen Arbeit des Übersetzers geworden. Diese Systeme basieren auf sogenannten *Translation Memories*: Sie sind das Werkzeug, das es seinen Benutzern ermöglicht, bereits übersetzte Texte zu speichern und sie für eine mögliche Wiederverwendung zu konsultieren. Es werden die Funktionsweise, der Aufbau und die Architektur von CAT-Systemen erläutert, wobei die wichtigsten Unterschiede zur maschinellen Übersetzung hervorgehoben werden. Sowohl die CAT-Systeme als auch die maschinellen Übersetzungssysteme sind weit verbreitet, und die Verwendung eines der Systeme schließt das andere nicht aus: In vielen Fällen verwenden auch Post-Editoren CAT-Tools in ihrem Arbeitsablauf. Zwei der wichtigsten Technologien, die professionelle Übersetzer benutzen, um ihre Produktivität zu steigern, sind nämlich maschinelle Übersetzungssysteme und *CAT-Tools*. *Translation Memories* können auch in maschinelle Übersetzungssysteme integriert werden und in dieser Dissertation wird beschrieben, wie das möglich ist.

Im zweiten Kapitel werden im Detail die verschiedenen Besonderheiten des Post-Editing von maschinellen Übersetzungen behandelt. Wie bereits erwähnt, ist das Post-Editing die Überarbeitung eines maschinell übersetzten Texts. Es gibt mehrere Unterschiede zwischen Post-Editing und Revision: Die Revision ist nämlich die Überarbeitung eines Texts, der von einem menschlichen Übersetzer produziert wurde. Außerdem macht ein Mensch nur selten lexikalische, syntaktische und grammatikalische Fehler, die hingegen in der Ausgabe einer maschinellen Übersetzung vorkommen können. Das Post-Editing der maschinellen Übersetzung wird durch eine internationale Norm, die ISO 18587, geregelt, deren Ziel es ist, Voraussetzungen an den Post-Editing-Prozess zu stellen. Es gibt auch wichtige Richtlinien für die Durchführung des Post-Editings, die von der Gesellschaft TAUS (Translation Automation User Society) zur Verfügung gestellt werden: Es gibt nämlich zwei Niveaus des Post-Editings, d. h. leichtes Post-Editing (*light PE*) und vollständiges Post-Editing (*full PE*). Ein leichtes Post-Editing soll nur eine verständliche Übersetzung ergeben, die daher nicht perfekt ist, während es mit einem vollständigen Post-Editing möglich ist, eine „Humanqualität“ zu erreichen. Die Kunden können ein bestimmtes PE-Niveau verlangen, an dem sich der Übersetzer festhalten muss, um seine Produktivität und Einnahmen nicht zu beeinträchtigen. Außerdem ist Post-Editing in der Regel zweisprachig: Dies geschieht beim Vergleich der beiden Texte, *source* und *target*, in den jeweiligen Sprachen. Trotzdem kann es auch einsprachig sein, wenn man nur den Zieltext betrachtet.

Jeder Übersetzer muss über verschiedene Kenntnisse verfügen, um seine Arbeit bestmöglich zu erledigen. Natürlich gilt das auch für Post-Editoren, die eine spezifische Ausbildung im Rahmen der Überprüfung der MÜ machen müssen. Zu den Fähigkeiten von Post-Editoren gehören beispielsweise das Erkennen von Fehlern, die Bewertung der Qualität des Zieltextes und die Verbesserung des Zieltextes. Obwohl die neuen neuronalen maschinellen Übersetzungssysteme viel genauer als die früheren SMÜ- und RBMÜ-Systeme sind, machen sie immer noch viele Fehler, die ein menschlicher Übersetzer normalerweise nicht machen würde. Die SMÜ-Systeme begehen systematische Fehler, die ganz einfach erkannt werden können, wie orthographische, grammatikalische oder syntaktische Fehler. Im Gegensatz dazu ist der Output von

NMÜ-Systemen flüssig, enthält aber oft Auslassungen oder andere Fehler, weswegen er nicht als perfekt bezeichnet werden kann. Zwei Schlüsselbegriffe stehen im Zusammenhang mit den Fehlern, die neuronale maschinelle Übersetzungssysteme verursachen: Qualität und Produktivität (d. h. die Anzahl der Wörter, die der Übersetzer im Laufe eines Tages nachbearbeiten kann). Diese beiden Begriffe sind eng miteinander verbunden, denn je besser die Qualität des Outputs, desto geringer ist die Verbesserungen des Post-Editors und desto höher ist seine Produktivität. Das führt zu einer Reduzierung der erforderlichen Zeit und dementsprechenden Kosten. Die Produktivität hängt jedoch von einer Reihe von anderen Faktoren ab, z. B. von der Sprachkombination, von der Art des Inhalts und von der Art des benutzenden MÜ-Systems.

Neben einer angemessenen Aus- und Weiterbildung in diesem Bereich muss der Post-Editor auch die Risiken kennen, die mit dieser Tätigkeit verbunden sind. In der Tat ist die Übersetzung eine Arbeit, die in all ihren Aspekten genau ausgeführt werden muss, weil sie selbst als "riskanter" Prozess mit rechtlichen Konsequenzen angesehen werden kann. Besondere Aufmerksamkeit muss dem Schutz der in die MÜ-Systeme eingegebenen Kunden- und Nutzerdaten gewidmet werden: Nicht alle MÜ-Systeme versichern den Benutzern einen totalen Datenschutz. Aus diesem Grund werden in dieser Arbeit die Richtlinien für die Verarbeitung personenbezogener Daten sowie die Risiken im Zusammenhang mit möglichen Schäden für den Kunden, IT-Risiken und Haftungsfragen sowie der Reputationsverlust des Buchenden selbst im Falle von Fehlern in der bestellten und gelieferten Arbeit analysiert.

Ein Abschnitt der Arbeit wurde der Zukunft des Post-Editing gewidmet: Einerseits gibt es eine unter den Übersetzern allgemein verbreitete Befürchtung, dass sie in den Hintergrund gedrängt werden und die klassische Übersetzungsarbeit in Zukunft nicht mehr existieren und nur noch Platz für maschinelle Übersetzung und Post-Editing bleiben wird; andererseits besteht die Hoffnung auf eine andauernde Verbesserung der maschinellen Übersetzung, um die sogenannte Singularität zu erreichen, bei der ein professioneller menschlicher Übersetzer weniger Zeit für Post-Editing als für die Revision einer (menschlichen) Übersetzung benötigt. Es ist bekannt, dass künstliche Intelligenz einen großen Einfluss auf die Arbeitswelt, einschließlich des Übersetzungsmarktes, haben wird. Dies wird zu einer Änderung der Aufgaben des

Übersetzers führen: Die Figur des Post-Editors selbst kann tatsächlich als eine Weiterentwicklung der Rolle des Übersetzers interpretiert werden.

Schließlich wurden die Hauptgründe analysiert, aus denen sich Übersetzer für oder gegen das Post-Editing entscheiden: Es gibt nämlich viele Übersetzer, die das Post-Editing immer noch ablehnen, obwohl es auf dem Übersetzungsmarkt und bei den Kunden sehr gefragt ist. Viele Übersetzer halten Post-Editing für langweilige und unproduktive Arbeit. Aber einer der Hauptgründe, weshalb sich viele Übersetzer nicht mit Post-Editing beschäftigen möchten, ist das Fehlen einer Zahlungsmethode, die den Aufwand für das Post-Editing berücksichtigt.

Das dritte und letzte Kapitel befasst sich mit dem zentralen Thema der Arbeit, nämlich der Vergütung von Post-Editing. Gegenwärtig gibt es keine Vergütungsmethode, die für alle Post-Editoren und Kunden gültig ist und von allen benutzt wird. In der Regel wird die Zahlungsmethode für Post-Editing von den Kunden vorgegeben. Es gibt verschiedene Arten der Vergütung, die am häufigsten verwendeten und bekanntesten sind jedoch die Bezahlung auf Wortbasis, auf Stundenbasis und nach Editierdistanz (*edit-distance*). Für jede der genannten Methoden gibt es Vor- und Nachteile. Die Grundlage des Wortbasismodells ist die herkömmliche Analyse der Vergütung von *CAT-Tools*, und ist auch am weitesten verbreitet: Dieses Modell ist für den Kunden vorteilhaft, weil es ihm Sicherheit über die Kosten des Auftrags gibt; wenn die Qualität des Outputs aber sehr schlecht ist, ist dieses Modell für den Post-Editor nachteilig. Die Bezahlung nach Editierdistanz misst die minimale Anzahl von notwendigen Operationen, um einen Satz in einen anderen zu ändern. Viele Kunden verwenden dieses Modell, um die Bemühung des Post-Editors zu vergüten, was jedoch nicht möglich ist, da die Editierdistanz nur die Anzahl der Operationen darstellt. Schließlich ist die Vergütung auf Stundenbasis das bevorzugte Modell der meisten Übersetzer, da sie die Qualität des MÜ-Outputs berücksichtigt. Es wird jedoch nicht oft verwendet.

Seit Jahren wird darüber diskutiert, welche dieser Methoden als die beste angesehen werden kann, d. h. ein Bezahlmodell, das mehrere Variablen berücksichtigt, wie z. B. die Qualität des vom maschinellen Übersetzungssystem erzeugten Outputs, das Sprachenpaar, den Inhalt des Textes und die Zeit, die für das Lesen des Ausgangstextes

und des Outputs notwendig ist. Bis heute gibt es keine endgültige Antwort auf diese Frage: Die meisten Post-Editoren werden normalerweise auf Wortbasis bezahlt. Das Wortbasismodell ist auch die Methode, die in der Regel in der Übersetzung verwendet wird. Aber, wie bereits erwähnt, wird die Bezahlung auf Wortbasis jedoch nicht immer als geeignet für Post-Editing angesehen, da die oben genannten Variablen nicht berücksichtigt werden können: Unter diesem Gesichtspunkt wäre eine Bezahlung auf Stundenbasis angemessener. Das Ziel dieses Kapitels ist es, die Meinungen der Post-Editoren zu diesem Thema darzustellen. Dazu wurde ein Fragebogen an 132 Teilnehmer weitergeleitet, die in zwei Gruppen von jeweils 66 Personen unterteilt wurden: eine Gruppe mit nur Post-Editoren, und eine mit dem allgemeinen Publikum. Daher nahmen auch Personen an der Studie teil, die keine Experten auf diesem Gebiet waren. Es wurde keine Altersgrenze festgelegt, weshalb die Teilnehmer eine Altersgruppe von 18 bis über 65 Jahren repräsentieren. Es gab keine Einschränkungen hinsichtlich des Geschlechts, des Bildungsabschlusses oder der geografischen Herkunft: Der Fragebogen wurde in englischer Sprache abgefasst, um den sprachlichen Bedürfnissen der verschiedenen Teilnehmer gerecht zu werden. Zahlreiche Verbände von Übersetzern und Dolmetschern mit Sitz in verschiedenen Teilen der Welt wurden kontaktiert und einige ihrer Mitglieder nahmen daran teil. Das Ziel war es auch, die größte Anzahl an Wahrnehmungen und Ideen der einzelnen Teilnehmer zum Thema zu sammeln. Was das allgemeine Publikum betrifft, so wurde der Fragebogen an Bekannte, Freunde und Verwandte weitergeleitet und auf einigen sozialen Kanälen (z. B. WhatsApp und Instagram) geteilt.

Der Fragebogen ist in zwei Abschnitte unterteilt: Der erste Abschnitt für alle 132 Teilnehmer befasst sich mit dem Verständnis der Wahrnehmung von Post-Editing, der Analyse des Wissens über diese Arbeit und der bestehenden Vergütungsmethoden; der zweite Abschnitt wurde jedoch nur von Post-Editoren ausgefüllt, da ihre persönlichen Meinungen zu Zahlungsmethoden untersucht werden sollten, um zu verstehen, wie ein gerechtere Vergütungssystem für Post-Editing aussehen könnte.

Um die Wahrnehmungen der Teilnehmer besser zu verstehen, wurde Letztere im ersten Teil des Fragebogens eine Aufgabe freier Assoziation gestellt: Es wurden zwei Stimulus-Wörter, *translation* (Übersetzung) und *Post-Editing*, vorgeschlagen. Die Teilnehmer mussten dann zwischen zwei und fünf Substantive oder Adjektive schreiben, die ihrer Meinung nach mit den zwei Stimulus-Wörtern verbunden waren.

Anschließend wurden eine Reihe geschlossener und offener Fragen zum Wissen über das Post-Editing und dessen Vergütungsmodalitäten gestellt.

Im zweiten Teil wurde eine Reihe von geschlossenen Fragen mit einfacher oder mehrfacher Auswahl vorgeschlagen, um herauszufinden, wie zufrieden die Post-Editoren mit der Art der Bezahlung sind, die sie normalerweise erhalten, und was ihrer Meinung nach die gerechteste Bezahlung sein könnte.

Abschließend wurde ein letzter kleiner Abschnitt den persönlichen Daten (Alter, Geschlecht und Studientitel) gewidmet, in dem die Teilnehmer einige individuelle Ansichten zu diesem Thema äußern konnten: Es wurden mehrere Aussagen von Post-Editoren gesammelt, die dazu beitragen, wichtige Denkanstöße zu geben.

Am Ende der Forschungsarbeiten konnten die Ausgangshypothesen bestätigt werden. Eingangs gab es zwei Voraussetzungen: Es wurde erwartet, dass in der Öffentlichkeit ein Mangel an Wissen über das Post-Editing und seine Vergütung besteht, sowie eine allgemeine Unzufriedenheit der Post-Editoren mit ihrem Verdienst und ein daraus resultierender Wunsch nach einer gerechteren Bezahlung. Es wurde daher angenommen, dass Post-Editing nicht angemessen vergütet wird und dass Post-Editoren nicht genug verdienen. Darüber hinaus gab es eine dritte Erwartung, nämlich dass die Mehrheit der teilnehmenden Post-Editoren normalerweise auf Wortbasis vergütet wird. Ziel der Studie war es daher auch herauszufinden, ob eine Bezahlung auf Stundenbasis eine gerechtere Vergütungsmethode sein könnte.

Wie oben dargestellt, besteht der Fragebogen hauptsächlich aus zwei Teilen. Im ersten Teil war die Aufgabe freier Assoziation sehr wichtig, weil sie es ermöglichte, die Ansichten der Teilnehmer zu diesem Thema zu bestimmen und die Unterschiede zwischen der Wahrnehmung dieser Arbeit und jener der Übersetzung hervorzuheben. Die Mehrzahl der Begriffe, die sich aus dem Stimulus-Wort „Übersetzung“ (*translation*) ergeben, haben im Allgemeinen eine positive Konnotation und weisen darüber hinaus eine größere Einheitlichkeit in den erfassten Antworten auf. Im Gegensatz dazu zeigen die Begriffe im Zusammenhang mit dem Stimulus-Wort „Post-Editing“ die Zweiteilung, die in dieser Welt besteht. Die Begriffe im Zusammenhang mit dem Post-Editing können in zwei große Kategorien unterteilt werden: Einerseits gibt es die Wörter, die ein positives Bild von diesem Beruf vermitteln und andererseits solche, die eine sehr

negative Konnotation haben. Die Übersetzung wird von den Teilnehmern als ein kreativer Prozess definiert und als eine verschiedene Sprachen und Kulturen miteinander verbindende Brücke beschrieben. Das Übersetzen wird aber auch als eine intensive und präzise Arbeit beschrieben, die auch eine Reihe von Schwierigkeiten aufweist. Aus diesem Grund können nur Menschen mit den angemessenen Kenntnissen und Fähigkeiten dank einer Ausbildung eine Übersetzung anfertigen. Darüber hinaus wird die Übersetzung als ein interessanter, nützlicher, unverzichtbarer und in manchen Fällen auch lustiger Beruf wahrgenommen. Für viele Teilnehmer bedeutet das Wort "Übersetzung" auch Exzellenz, Professionalität, Vertrauen, Leidenschaft und Freiheit; diese Arbeit ermöglicht auch eine angemessene Bezahlung. Einige beschreiben diesen Beruf als zeitaufwendig, während andere sagen, dass er unter bestimmten Umständen sogar schneller als Post-Editing ist. Schließlich äußerten einige Teilnehmer ihre Besorgnis darüber, dass die (menschliche) Übersetzung heutzutage vernachlässigt wird und droht, durch maschinelle Übersetzung überholt zu werden.

Bei den Begriffen, die sich aus dem Stimulus-Wort "Post-Editing" ergaben, sind die Ergebnisse jedoch weniger einheitlich: In diesem Fall wurden häufiger Substantive und Adjektive genannt, die beschreiben, worin diese Arbeit besteht, z. B. Wörter wie *revision*, *correction*, *adjustment*, usw. (d. h. Überarbeitung, Korrektur, Anpassung). Diese Wörter drücken in der Tat aus, dass es sich um die Revision der Ausgabe eines maschinellen Übersetzungssystems handelt, die durch die Korrektur der vorhandenen Fehler erfolgt. Um die tatsächliche Wahrnehmung dieses Berufs zu verstehen, sind jedoch jene Begriffe interessanter, die zwar weniger häufig genannt wurden, aber die persönliche Meinung der einzelnen Teilnehmer wiedergeben: Insbesondere können die Post-Editoren selbst einen wichtigen Einblick in das Phänomen geben, da es sich um ihren eigenen Beruf handelt. Wie bereits erwähnt, unterscheiden sich die zitierten Aussagen stark voneinander und zeigen die großen Unterschiede in der Wahrnehmung des Post-Editing, insbesondere unter den Post-Editoren selbst. In der Tat gibt es Anhänger dieses Berufs, die ihn als modern und schnell beschreiben, da er Übersetzungszeit spart und eine einfache Korrektur von Fehlern ermöglicht. Einige glauben auch, dass Post-Editing interessant, innovativ, effizient und sogar unverzichtbar ist. Zu den Gegnern gehören diejenigen, die Post-Editing-Arbeit als äußerst langweilig empfinden, weil sie repetitiv und un kreativ erscheint. Viele Übersetzer sind der

Meinung, dass die Qualität der maschinell erstellten Übersetzung zu schlecht ist und dass die Korrektur viel Zeit in Anspruch nimmt. Das Post-Editing wird als schwierige und frustrierende Aufgabe empfunden, die voller Hindernisse und den Aufwand nicht wert ist. Einige Teilnehmer behaupten, dass es sich um eine oberflächliche und ungeeignete Arbeit handle, da sie nicht als echte Übersetzung angesehen werden kann, und dass, wie ein Übersetzer betonte, sich die Post-Editoren wie "Sklaven" fühlten (der betreffende Teilnehmer benutzte das Wort *slave*, um Post-Editing zu beschreiben). Darüber hinaus gibt es zwei Denkrichtungen in Bezug auf die Vergütung: Auf der einen Seite gibt es diejenigen, die denken, dass man mit Post-Editing Kosten reduzieren kann; auf der anderen Seite gibt es diejenigen, die behaupten, dass Post-Editing wegen der bestehenden Kosteneinsparungen extrem unterbezahlt ist.

Anschließend wurden die Antworten auf die offenen und geschlossenen Fragen analysiert. Obwohl mehrere Personen aus der breiten Öffentlichkeit angaben, vor der Beantwortung des Fragebogens von Post-Editing gehört zu haben, war fast niemand in der Lage, die bestehenden Vergütungsmodalitäten für diese Arbeit genau zu nennen. Die Teilnehmer aus der breiten Öffentlichkeit könnten in der Tat mögliche gegenwärtige oder künftige Kunden darstellen: Aus diesem Grund wird ein geringes Wissen über Post-Editing und dessen Vergütung nicht als positiver Faktor angesehen. Laut 88 % der Post-Editoren, die an der Studie teilgenommen haben, haben die Öffentlichkeit und folglich auch die Kunden keine richtige Vorstellung von dieser Arbeit und ihrer Vergütung. Dieser Mangel an Information führt häufig zur Anwendung eines unangemessenen und ungeeigneten Vergütungsmodells. Dies führt zu Arbeitsbedingungen, die nur für die Auftraggeber selbst vorteilhaft sind. Das Hauptziel der Kunden besteht in der Kostensenkung, indem die normalerweise für die Übersetzung benötigte Zeit reduziert wird. Die Auftraggeber wollen aber, dass die Qualität des Ergebnisses mit der menschlichen vergleichbar ist.

Die Post-Editoren wurden gefragt, ob sie mit ihrer Arbeit und ihrem Einkommen zufrieden seien. Die Analyse der Antworten zeigt, dass viele von ihnen unzufrieden sind. Tatsächlich gab fast die Hälfte der Befragten an, dass sie mit dem Vergütungssystem, das sie erhalten, unzufrieden sind. Erwartungsgemäß wird die Mehrheit der Post-Editoren auf Wortbasis bezahlt, aber viele würden ein gerechteres Vergütungssystem bevorzugen: In der Tat berücksichtigt die Bezahlung auf Wortbasis

sehr oft nicht die Komplexität der Arbeit von Post-Editing. Post-Editing besteht nicht nur aus der einfachen Korrektur von Fehlern in der Ausgabe, sondern aus einer langen Reihe von Aufgaben, wie dem Lesen der Ausgabe und des Ausgangstextes, der Identifizierung der Fehler und deren Korrektur. Der endgültige Zieltext muss noch einmal aufmerksam durchgelesen und dann dem Kunden geliefert werden. Obwohl sich der von neuronalen Übersetzungssystemen erzeugte Output in den letzten Jahren deutlich verbessert hat, muss er noch weiter verfeinert werden. In der Tat ist es noch nicht möglich, insbesondere bei bestimmten Sprachpaaren eine Qualität zu erreichen, die der menschlichen Qualität entspricht oder sehr ähnlich ist. Die schlechte Qualität des Outputs ist eines der Hauptprobleme von Post-Editing und macht diesen Überarbeitungsprozess wesentlich zeitaufwendiger und komplexer. Die kognitive Anstrengung ist dann größer. In manchen Fällen bevorzugen die Post-Editoren, einen ganzen Abschnitt neu zu übersetzen, anstatt ihn zu überprüfen, da die Qualität niedrig ist und zu viel Zeit für das Post-Editing aufgewendet werden müsste: Dies hat einen starken Einfluss auf die Produktivität der Post-Editoren und folglich auch auf ihren Verdienst.

Die verwendete Zeit und die Qualität des MÜ-Outputs sind also zwei grundlegende Variablen, die jedoch bei der Bezahlung auf Wortbasis normalerweise nicht ausreichend berücksichtigt werden. Wie einige Teilnehmer bemerkten, wäre eine Bezahlung auf Stundenbasis vorzuziehen, die es den Post-Editoren ermöglicht, nicht unter Druck zu arbeiten und die verschiedenen Elemente, die ihre Arbeit beeinflussen, wirklich zu berücksichtigen. Nicht selten geben Übersetzer an, dass sie sich beim Post-Editing unter Druck fühlen, weil sie die Arbeit in möglichst kurzer Zeit erledigen müssen, um negative Auswirkungen auf ihre Einnahme zu vermeiden. Die endgültige Qualität muss aber, wie bereits erwähnt, der Humanqualität entsprechen. Es sollte auch darauf hingewiesen werden, dass die Nachbearbeitung der maschinellen Übersetzung nicht für alle Textarten geeignet ist. Die Kunden berücksichtigen diesen wichtigen Faktor jedoch nicht: Das führt zu einer sehr schlechten Qualität der MÜ-Ausgabe, da der Inhalt des Ausgangstextes nicht für die Übersetzung durch ein maschinelles Übersetzungssystem geeignet ist. Dennoch gibt es heutzutage auf dem Übersetzungsmarkt eine steigende Nachfrage nach Post-Editing und bei der Analyse der Gründe, warum sich die Post-Editoren für diese Arbeit entschieden haben, gab die Mehrheit der Teilnehmer an, dass

es sich um eine unvermeidliche Wahl handle, da es eine der gefragtesten Tätigkeiten auf dem Zielmarkt sei. Einige Post-Editoren betonten, dass sie in vielen Fällen nur durch diese „obligatorische“ Wahl den Job bekommen würden.

Die Post-Editoren gefragt, was ihrer Meinung nach die angemessenere Zahlungsmethode für Post-Editing sei. Aus der Analyse der Antworten geht hervor, dass 44.1 % der Teilnehmer eine Bezahlung auf Stundenbasis bevorzugen würden, da diese das gerechteste Bezahlungsmodell für Post-Editing sei. Ein beträchtlicher Teil der Post-Editoren bevorzugt jedoch eine Bezahlung auf Wortbasis (33.8 % der Post-Editoren). Die Bezahlung auf Wortbasis wird nämlich nicht immer auf dieselbe Weise berechnet, da einige Auftraggeber auch die Editierdistanz berücksichtigen oder sogar nicht nur die zu bearbeitenden Wörter, sondern auch die Wörter im Ausgangstext bezahlen, wie z. B. das Unternehmen „Translated“. In vielen anderen Fällen entspricht die erhaltene Vergütung jedoch nicht den Erwartungen der Post-Editoren. Viele sind der Meinung, dass nur eine Bezahlung auf Stundenbasis wirklich alle Merkmale des Textes sowie die tatsächliche kognitive Anstrengung berücksichtigen kann, die das Post-Editing erfordert.

Der Mangel an Information der Kunden repräsentiert den Grund dieser seit mehreren Jahren geführten Diskussion. Insbesondere sollten genauere Richtlinien für die akzeptierten Zahlungsmethoden festgelegt werden: Post-Editing ist eine Tätigkeit, die seit mehreren Jahren besteht und nicht nur ihre Schwierigkeiten, also ihre Schwächen, sondern auch ihre Stärken allgemein bekannt sind. Post-Editing ermöglicht eine echte Produktivitätssteigerung und die Reduzierung des Zeitaufwands, wenn es korrekt und mit den richtigen Werkzeugen durchgeführt wird. Es ist jedoch notwendig, die Übersetzer, die diese Tätigkeit ausüben, besser zu schützen, da es auf dem Übersetzungsmarkt selbst eine große Nachfrage nach Post-Editing besteht. Auf diese Weise könnte eine einheitliche und gerechte Zahlungsmethode eingeführt werden, die für alle Post-Editoren und alle Kunden gilt. Gegenwärtig herrscht hingegen ein allgemeiner Informationsmangel der Kunden, die in den meisten Fällen über die anzuwendende Vergütungsmethode und die Bedingungen für ihre Anwendung entscheiden müssen. Post-Editing könnte auf diese Weise mehr Aufmerksamkeit und Anerkennung erhalten. Darüber hinaus sollten genauere Angaben zu den Textsorten und

zu den Inhalten bereitgestellt werden, die durch Post-Editing überarbeitet werden können, worüber die Kunden informiert werden sollten. Wenn möglich, sollten Post-Editoren wiederum eine zu niedrige Bezahlung ablehnen, damit der Kunde den Wert ihrer Fähigkeiten und ihrer Arbeit erkennt. Wenn die Zukunft der Übersetzung im Post-Editing liegt, ist es notwendig, eine angemessene Vergütung zu erreichen, um neue zukünftige Übersetzer (und Post-Editoren) für die Welt der Übersetzung zu gewinnen und sie nicht zu vertreiben, weil sie durch die vorgeschlagenen Arbeitsbedingungen entmutigt werden.