

Il Multidimensional Quality Metrics nell'analisi qualitativa della traduzione automatica

Gennaro Nolano, Johanna Monti

UNIOR NLP Research Group

Università degli Studi di Napoli L'Orientale

Il seguente articolo è un estratto di un lavoro più ampio sulla traduzione automatica del testo King, Warrior, Magician, Lover di Robert Moore e Douglas Gillette. Nel corso del suddetto lavoro questo testo, dopo essere stato analizzato linguisticamente, è stato tradotto attraverso i due traduttori automatici Google Translate e DeepL per poi essere completato in fase di post-editing da traduttore umano. Fase essenziale del processo è stata sicuramente la valutazione degli errori riscontrati nelle due traduzioni automatiche, effettuata con l'aiuto del Multidimensional Quality Metrics (di qui in avanti indicato con la sigla MQM), una metrica di valutazione per traduzioni e testi più in generale. Nell'articolo qui presente si definirà questa metrica e il modo in cui è stata utilizzata nel lavoro summenzionato.

1 MQM

MQM (Harris et al., 2015) offre

a framework for describing and defining quality metrics used to assess the quality of translated texts and to identify specific issues in those texts. It provides a systematic framework to describe quality metrics based on the identification of textual features.¹ (...)

ed è formato da:

¹Un quadro per la descrizione e la definizione di parametri per la qualità, utilizzato per valutare la qualità dei testi tradotti e per identificarne problematiche specifiche. Offre un quadro sistematico per descrivere parametri di qualità basato sull'identificazione di caratteristiche testuali. (Traduzione mia)

- un vocabolario per la categorizzazione delle problematiche relative alla qualità;
- un meccanismo di assegnazione di un punteggio basato sul conteggio degli errori, o sulla loro annotazione;
- una normativa di markup, che include:
 - un meccanismo XML formale per descrivere i parametri di qualità;
 - diversi attributi del namespace [MQM:]² che possono essere usati sia con XML che con HTML5, progettati per funzionare con il Tag Set Internazionale 2.0 (ITS 2.0);
 - elementi del namespace MQM: che possono essere utilizzati per inserire dati MQM in un file XML nel caso elementi esistenti non soddisfino le necessità dell'utente).
- Una serie di linee guida per la selezione di problematiche, basate sulle specifiche ASTM F2575:2014³

La base teorica su cui si basa tale modello è di stampo assolutamente funzionalista, per cui il primo valore che va a valutare è la vicinanza del testo tradotto ai fini comunicativi proposti dall'originale. Dal documento informativo:

MQM is useful for assessing verifiable qualities of translations. It is not intended to

²Per namespace (o spazio dei nomi) si intende, in informatica, la definizione di un contesto per permettere al software di riconoscere ciò su cui deve lavorare. Vengono utilizzati in particolare in XML e HTML, e seguono la sintassi `NAMESPACE:IDENTIFICATORE`. Per ulteriori informazioni si rimanda a Bray et al. (2009)

³Per un'analisi più approfondita è possibile consultare l'indirizzo <https://www.gauchatranslations.com/wp-content/uploads/2017/06/ASTM-General-rv-2016-03-21.pdf>

address purely subjective criteria (such as “artistry” or “elegance”) that may be of key importance in some circumstances. Rather, it provides a functional approach to quality that seeks to see whether a translation meets specifications and to identify aspects that may fall short of expectations.⁴

Questo significa quindi che esso da solo non basta a verificare la qualità di determinati testi come pubblicitari, divulgativi o letterari, per i quali un certo livello di soggettività nel controllo della qualità non è solo utile, ma necessario e richiesto. Come abbiamo già visto in precedenza, la tipologia testuale e l'ambito di studio di *King, Warrior, Magician, Lover* fanno sì che nel testo l'autore non sia completamente invisibile, ma che, anzi, la sua presenza è molto spesso parte integrante del significato del libro, e per tale motivo il quadro MQM non poteva essere l'unico parametro utilizzato, ma, ciò nonostante, si è rivelato uno strumento estremamente utile per valutare il lavoro effettuato dai traduttori automatici presi in esame.

Termini utilizzati in MQM

L'analisi di qualità in MQM passa attraverso l'utilizzo di una terminologia particolare, consta dei seguenti termini, alcuni dei quali riprendono a piene mani dalla teoria della traduzione:

Accuratezza: il grado di conformità del contenuto del testo d'arrivo al contenuto del testo di partenza;

Adeguatezza: sinonimo di accuratezza, solitamente usato per la valutazione di testi tradotti con traduttori automatici;

Categoria di dati: termine teorico utilizzato per definire un particolare tipo di informazione (come peso, tipologia di problematiche e simili), utile a descrivere la qualità della traduzione;

Dimensione: un qualunque aspetto, relativo alla struttura superficiale di una traduzione, che rientri nel calcolo della valutazione della qualità. Le dimensioni utilizzate in MQM sono:

- Accuratezza;
- Compatibilità⁵;
- Convenzioni del locale;
- Fruibilità;
- Internazionalizzazione;
- Stile;
- Struttura;
- Terminologia;

⁴MQM è utile per valutare qualità quantificabili nelle traduzioni. Non è pensato per descrivere criteri puramente soggettivi (come ad esempio l'“abilità artistica” o l'“eleganza”) che potrebbero essere di importanza centrale in alcune circostanze. Piuttosto, offre un approccio funzionale alla qualità che cerca di valutare se una traduzione soddisfi i requisiti e identifichi aspetti che potrebbero non essere all'altezza delle aspettative. (Traduzione mia)

⁵Questa dimensione è da considerarsi obsoleta, e nel documento ne si sconsiglia l'utilizzo, conservandone la descrizione solo per motivi di compatibilità con versioni passate. Per questo motivo non verrà ulteriormente trattata in questa sede.

- Verità;
- Altro.

e verranno in seguito analizzate nel dettaglio;

Errore: viene definito errore una problematica la cui incorrettezza è stata verificata, in seguito a un primo controllo;

Gravità: indica il grado di gravità di una particolare istanza di una problematica. Il modello base di MQM utilizza quattro livelli di gravità: nulla, marginale, grave e critica;

Metodo: un modo di valutare alcuni aspetti della qualità di una traduzione;

Parametro: un aspetto della traduzione che definisce determinate aspettative del prodotto tradotto. Un esempio di parametro potrebbe essere il locale;

Penalità per l'errore: è il punteggio assegnato ad ogni errore in un testo secondo il sistema di misurazione analitico;

Peso: un valore numerico che indica quanto è importante una determinata problematica nella valutazione generale della qualità. Il peso di default è 1.0, e a peso maggiore corrisponde importanza maggiore, con il valore 0 che indica una problematica accertata ma non contata nel punteggio MQM. Il peso viene utilizzato come moltiplicatore delle penalità per errore;

Problematica: viene definita problematica un problema *potenziale* riscontrato nel testo. Ciò significa che una problematica potrebbe anche rivelarsi essere una scelta stilistica o linguistica, ma in principio reputata un possibile errore da un processo automatizzato o da un lettore umano;

Qualità: l'adesione del prodotto a specifiche particolari, che nel caso di testo tradotto sono un grado di accuratezza e di fruibilità relativi al pubblico di arrivi e ai fini del prodotto finale. Tale accuratezza e tale fruibilità entra in relazione con tutte le altre eventuali specifiche negoziate tra cliente e traduttore;

Sistema di misurazione: in MQM un sistema di misurazione definisce quanto un testo soddisfa i requisiti di qualità. Esso consta di una o più problematiche attraverso cui è valutato il testo, e un metodo di valutazione olistico o analitico;

Sistema di misurazione analitico: un sistema di misurazione che si basa sull'identificazione e la successiva classificazione di problematiche all'interno di un testo;

Sistema di misurazione olistico: un sistema di misurazione basato su questioni o dichiarazioni (accompagnate da definizioni che le rendono più chiare), corrispondenti a tipi di problematiche, utilizzate come base per la valutazione del testo, insieme a una scala di valori utilizzata per valutare i singoli elementi e standard di qualità per livelli specifici di prestazione;

Specifiche: una descrizione dei requisiti di una tradu-

zione. MQM utilizza un sottogruppo delle specifiche definite in ASTM F2575-14⁶.

1.1 Problematiche

MQM definisce più di 100 categorie di problematiche (Arle et al., 2015), ricavate da una disamina di sistemi di valutazione della qualità, sia basati su accertazione automatizzata delle problematiche, sia basati sul lavoro di revisori umani.

Tali categorie sono inserite in una gerarchia di specificazione che va dal generale al particolare. È quindi possibile, qualora richiesto dall'analisi, utilizzare solo uno specifico livello di gerarchia al fine di avere una certa granularità in fase di valutazione. In generale, nel documento originale si consiglia all'utente:

metrics should check the fewest number of issues possible to achieve the requirements of users.⁷ (...)

Al livello più alto (e quindi più generale) le problematiche vengono ricategorizzate in diverse dimensioni superiori, identificate da un nome (nel presente lavoro tradotto in italiano) e da un tag, definito nel documento originale come "valori ID" (nel presente lavoro lasciati in lingua inglese):

Accuratezza (*accuracy*): comprende problematiche relative al rapporto tra il testo d'origine e il testo di arrivo, tra cui figurano, ad esempio, eventuali slittamenti semantici e aggiunta/omissione di contenuto;

Convenzioni del locale (*locale-convention*): comprende problematiche relative al mancato rispetto di convenzioni relative al locale di arrivo (come ad esempio l'utilizzo di determinati simboli numerici, o traslitterazioni);

Fruibilità (*fluency*): comprende problematiche relative all'aspetto linguistico del prodotto (ovvero alla sua aderenza a un determinato codice linguistico) che, in quanto tali, possono essere applicate tanto a un testo tradotto quanto a un testo originale indipendentemente;

Internazionalizzazione (*internationalization*): comprende problematiche relative alla preparazione del prodotto di partenza per una futura traduzione;

Stile (*style*): comprende problematiche relative sia agli stili definiti formalmente (nelle guide di stile), sia a quelli definiti informalmente (come ad esempio potrebbe essere la richiesta di un cliente di avere un testo dallo "stile facile");

Design (*design*): comprende problematiche relative alla presentazione superficiale del prodotto tradotto;

Terminologia (*terminology*): comprende problematiche relative all'utilizzo di una terminologia specifica (non importa se essa sia specifica di un genere o specifica perché desiderata dal cliente);

Verità (*verity*): comprende problematiche relative all'idoneità del contenuto al locale e al pubblico di arrivo (un eventuale riferimento non necessario a una legge o a una festività anglosassone in un testo per un pubblico italiano rientrerebbe in questa dimensione);

Altro (*other*): comprende problematiche non definiti in nessuna delle dimensioni precedenti.

A queste dimensioni (definite in questo caso come "genitori" delle problematiche più specifiche) corrispondono poi le specifiche problematiche riscontrabili nel corso di una traduzione. Come già menzionato in precedenza, tali problematiche sono più di 100, ma per un utilizzo generale MQM descrive un "nucleo" (definito in originale *MQM core*) che comprende 20 tipologie di problematiche (tra cui rientrano le dimensioni descritte in precedenza) tra quelle più comuni nella verifica della qualità in un testo tradotto. Questa semplificazione, in generale indicata dalla guida MQM come è utile soprattutto a mantenere un alto livello di operabilità tra i sistemi, che è uno degli obiettivi principali dell'utilizzo di MQM.

Il nucleo di MQM è composto dalle seguenti tipologie di problematiche:

- Accuratezza(*accuracy*);

Aggiunta (*addition*): il prodotto di arrivo include porzioni non presenti nel prodotto di partenza;

Errori di traduzione (*mistranslation*): il contenuto del prodotto di arrivo non rappresenta accuratamente il contenuto del prodotto di partenza ;

Mancata traduzione (*untranslated*): parte del contenuto che sarebbe dovuto essere tradotto è stato lasciato in originale;

Omissione (*omission*): mancanza nel testo di arrivo di contenuto presente nel prodotto di partenza;

- Convenzioni del locale(*locale-convention*);

- Fruibilità(*fluency*);

Grammatica (*grammar*): problematiche relative alla grammatica o alla sintassi del testo;

Incoerenza (*inconsistency*): il testo presenta incoerenze interne;

Ortografia (*spelling*): problematiche relative all'ortografia delle parole;

Registro grammaticale (*grammatical-register*): problematiche relative a un'incongruenza tra il registro richiesto e quello utilizzato;

⁶In ASTM F2575-14, le specifiche riguardano i parametri della traduzione (ovvero le specifiche negoziate tra cliente e traduttori), le informazioni del contenuto di partenza, i requisiti del contenuto di arrivo, i parametri del processo di traduzione, l'ambiente di lavoro e le relazioni delle parti interessate con il progetto.

⁷[...] la metrica dovrebbe contenere il minor numero possibile di problematiche necessarie a venire incontro alle richieste degli utenti. (Traduzione mia)

Tipografia (typography): problematiche relative alla rappresentazione tecnica del prodotto;

Indefinito (unintelligible): problematica la cui esatta natura non è determinabile, ma che ciò nonostante causa gravi problemi nella fruibilità del testo;

- Internazionalizzazione (*internationalization*);
- Stile (*style*);
- Struttura (*design*);
- Terminologia (*terminology*);
- Verità (*verity*);

Completezza (completeness): il testo è incompleto;

Requisiti legali (legal-requirements): il testo non soddisfa i requisiti legali indicati nelle specificazioni;

Contenuto specifico del locale (locale-specific-content): problematiche relative al mantenimento di contenuto specifico al locale di partenza;

- Altro (*other*).

Estensioni dell'utente

L'utente ha la possibilità di definire proprie problematiche (benché gli stessi creatori consiglino di limitarsi alle problematiche pre-definite da MQM), le quali devono essere necessariamente definite dalle seguenti informazioni:

- Nome della problematica;
- Valore ID, un QName⁸, utilizzato come identificatore XML del tipo di problematica. L'ID deve essere formato da un nome locale (definito dall'utente) accompagnato dal prefisso x-, che ne indica l'appartenza alla lista di problematiche definite dall'utente. Un esempio potrebbe essere X-RESPEAKING-ERROR;
- Una problematica genitore, ovvero il valore ID di una tipologia di problematiche più generale che comprende quella definita dall'utente. La problematica genitore può essere sia una tipologia definita dall'utente che una tipologia pre-definita in MQM;
- Definizione, ovvero una spiegazione della tipologia di problematica.

1.2 Assegnazione del punteggio

MQM offre un metodo pre-definito per assegnare un punteggio alla qualità di una data traduzione, attraverso l'utilizzo dei seguenti valori:

Peso: il valore assegnato a ogni singola problematica nel calcolo della qualità del contenuto finale. Di

⁸Un QName, ovvero dall'inglese Qualified Name, è un nome qualificativo unico per indicare un elemento, un attributo o un identificatore in un documento XML. È formato da un nome locale a cui è affibbiato un prefisso, corrispondente al suo namespace.

default, il valore assegnato al peso è 1.0, e può essere aggiornato dall'utente a seconda di bisogni specifici;

Gravità: i livelli di gravità sono definiti attraverso un grado di gravità e un numero identificato come *moltiplicatore di problematica*, utilizzato della formula per il calcolo della qualità totale del contenuto. I livelli definiti sono i seguenti:

- **Nulla:0.** Problematiche con questa gravità si riferiscono a entità che devono essere evidenziate per poi venire in eseguito analizzate o risolte, ma che non hanno peso negativo per la traduzione. Questo livello di gravità può essere visto come un segnale di avviso che non comporta alcuna penalità, come possono essere "errori preferenziali" (ovvero entità non errate di per sé, ma per cui il revisore o il cliente potrebbe preferire una soluzione diversa) o errori sistematici ripetuti e facilmente risolvibili.
- **Marginale:1.** Problematiche con questa gravità non hanno alcun impatto sull'usabilità o la comprensibilità del contenuto. Sono errori di severità minore quelli facilmente risolvibili del lettore/utente senza che l'usabilità ne venga intaccata, e per tale motivo la loro risoluzione è a discrezione dei responsabili del contenuto. Ne è un esempio uno spazio aggiuntivo in seguito a un punto.
- **Grave:10.** Problematiche di questo tipo intaccano l'usabilità o la comprensibilità del contenuto, senza però renderlo inutilizzabile. Rientrano in questa categoria errori che non possono essere sempre corretti dal lettore/utente, ed è quindi generalmente consigliato correggerli prima dell'utilizzo del contenuto. Ne può essere un esempio un errore ortografico, che renderebbe difficile, ma non impossibile, comprendere il significato di una parola.
- **Critico:100** Problematiche critiche rendono il contenuto inadatto all'uso. Se l'errore non permette al lettore/utente di utilizzare il contenuto come pianificato, o se porta alla presenza di informazioni errate che possono causare danni all'utente, esso deve essere classificato come critico, e in quanto tale deve essere corretto prima dell'utilizzo del prodotto, poiché anche un solo errore critico potrebbe creare seri problemi. Ne è un esempio un grave errore grammaticale che modifica il significato nel testo.

Utilizzando i valori sopra descritti è possibile calcolare la penalità (P) del contenuto d'arrivo e di quello di partenza, attraverso la seguente formula:

$$P = (PM + PG * MG + PC * MC) / NP$$

Dove

- PM = numero di problematiche con gravità “marginale”;
- PG = numero di problematiche con gravità “grave”;
- PC = numero di problematiche con gravità “critica”;
- MG = moltiplicatore delle problematiche “gravi”;
- MC = moltiplicatore delle problematiche “critiche”;
- WC = numero di parole del contenuto.

Infine, è possibile calcolare la qualità finale del contenuto (TQ) con la formula

$$TQ = 100 - TP + SP$$

Dove

- TP (Target content penalties) = penalità per il testo di arrivo;
- SP (source content penalties) = penalità per il testo di partenza.

Come già menzionato in precedenza, ogni problematica ha, di default, un peso uguale a 1.0. Tale valore può essere però modificato a seconda delle necessità dell'utente, così che, ad esempio, un determinato tipo di errore possa avere un peso maggiore nel calcolo del punteggio della qualità del prodotto.

Si fa notare che la presenza delle penalità relative al testo di partenza nel calcolo della qualità del contenuto finale implica che, qualora il testo sorgente sia a disposizione del traduttore, è sua ulteriore responsabilità andare a risolvere eventuali problemi nati da un testo difettoso. Nel caso il testo di partenza non fosse disponibile, la sua penalità è di default calcolata come 0, così da non influire sul punteggio finale.

2 Utilizzo pratico dello standard MQM

2.1 Definizioni necessarie prima dell'utilizzo

Nel presente lavoro lo standard di valutazione MQM è stato principalmente utilizzato per valutare i risultati ottenuti dai traduttori automatici Google Traduttore e DeepL sul testo da noi preso in esame. In particolare, il segmento tradotto è stato una selezione dal libro intero, ovvero il capitolo *Introduzione* intero accompagnato da una sezione esemplificativa di ognuno dei capitoli dedicati ai quattro archetipi principali. In totale, quindi, il testo tradotto è definito da cinque capitoli diversi: *Introduzione, King, Warrior, Magician e Lover*.

La valutazione delle problematiche è stata effettuata attraverso la creazione di tabelle SQL strutturate nel seguente modo:

- Per colonne:

- Id identificativo del traduttore automatico preso in esame (GT per Google Traduttore e DL per DeepL);
 - Sezione estratta dal testo in cui si è verificata la problematica;
 - Tipologia di problematica secondo la lista indicata in precedenza;
 - Valore numerico del livello di gravità.
- Per righe le singole problematiche.

Notiamo innanzitutto che nella nostra analisi il peso delle problematiche è considerato sempre di default 1.0, così che abbiano tutte la stessa importanza nel calcolo finale. Inoltre, la penalità per il testo di partenza SP ha valore 0, in quanto il testo originale su cui abbiamo lavorato non presentava alcuna problematica. Una seconda nota indicativa riguarda la finalità della traduzione, il suo *skopos* (Vermeer e Reiß, 1984). Secondo la definizione summenzionata di MQM si nota subito che quest'unità di misura è legata all'“uso” che si farà del prodotto finale. Nel nostro caso, quest'uso può essere identificato con la completa comprensione, da parte di un pubblico più o meno identificato, del testo in ogni sua parte. Una mancata comprensione potrebbe nascere da situazioni diverse: prima di tutto, potrebbe essere causata da un'errata o mancata traduzione della terminologia. A queste problematiche è stata assegnata gravità di valore 10 o 100, seppur facilmente risolvibili aggiungendo o modificando il glossario dei traduttori automatici.

Una seconda causa di mancata comprensione del testo potrebbe essere invece un errore ‘grossolano’, e con questo termine definiamo di visibilità elevata, ma la cui risoluzione è impossibile senza avere il testo originale a propria disposizione (questi errori rientrano molto spesso nella categoria *mistranslation* di elevata letteralità della traduzione, con una gravità ovviamente di 100).

Una terza causa potrebbe essere la mancata aderenza alle aspettative del pubblico, per cui si lascia per implicito un'informazione che deve invece essere esplicita perché si possa comprendere un passaggio. Per evidenziare problemi di questo tipo, è necessario prima di tutto identificare le caratteristiche del teorico pubblico di arrivo. In linea puramente astratta, quindi, definiremo il lettore medio della nostra traduzione come:

- un madrelingua italiano;
- senza un alto livello di conoscenza della lingua inglese;
- senza un alto livello di conoscenza della cultura di partenza (nel nostro caso quella statunitense);
- senza un alto livello di conoscenza dell'ambito di studio del testo né della relativa tipologia testuale.

Il primo punto è risolvibile accertandosi che il testo utilizzi un linguaggio coerente a quello che verrebbe utilizzato se fosse un testo originale in lingua italiana, anche se, come abbiamo notato in precedenza, molto

spesso in caso di generi testuali senza una forte tradizione nella lingua/cultura d'arrivo, si potrebbe tendere allo straniamento anche del linguaggio utilizzato (o della strutturazione del testo). I tre punti rimanenti possono essere risolti in diversi modi, ad esempio, attraverso il mantenimento dell'espressione originale con l'eventuale aggiunta di una nota (a piè pagina o inclusa nel testo, segnalata o meno), oppure attraverso, laddove sia possibile, di tecniche di addomesticamento. Queste risoluzioni, in ogni caso, sono risolvibili, in ambito lavorativo, solo da un chiarimento da parte del cliente. Ovviamente, laddove il traduttore fosse libero di agire liberamente, è egli stesso a decidere come risolvere la cosa.

Abbiamo inoltre menzionato a più riprese la necessità di chiarire con il cliente, in fase di pre-trasferimento, come risolvere determinati punti critici nel testo. Questo chiarimento diventa assolutamente necessario nel momento di definizione di cosa dev'essere calcolata come problematica e di quanto questo influisca sulle aspettative del prodotto finale, e per questo sarà necessario in questa sezione discuterne.

Abbiamo già definito questi punti critici risolvibili solo attraverso una disambiguazione da parte del cliente come suddivisibili in due aree principali:

- Problematiche tecniche;
- Problematiche etiche.

Nel primo gruppo rientravano le seguenti questioni, che indicheremo assieme alla soluzione che abbiamo deciso di utilizzare (tenendo ben presente che non vuole essere una soluzione oggettiva, ma semplicemente ciò che è stato scelto in questa sede):

Problematica	Soluzione
Riferimenti a fatti e concetti poco o per nulla noti dal pubblico d'arrivo, o persino assenti nella cultura d'arrivo	Mantenimento della forma originale con aggiunta di note a piè pagina
Riferimento al lettore	Tradotto con forme impersonali più formali
Utilizzo di termini in lingua inglese laddove l'italiano può scegliere tra questi e la forma italiana	Scelta la forma italiana
Utilizzo del trattino per segnalare le incidentali	Mantenuto anche in italiano
Traslitterazioni dei nomi russi	Utilizzo della traslitterazione scientifica piuttosto che quella anglosassone
Segnalazione di termini tecnici	Utilizzo del corsivo

Tabella 1: *Risoluzione di problematiche di tipo tecnico*

Altre problematiche potrebbero riferirsi anche al carattere utilizzato, alla struttura del testo o alla sua formattazione, fatto eventualmente riguardante non solo il testo cartaceo, ma anche quello digitale, spesso segnato da limitazioni di diverso tipo (come potrebbe essere, ad esempio, il peso del file).

Nel secondo gruppo rientravano invece problematiche più difficili da risolvere, riguardanti il punto di vista del testo. Queste problematiche nascono, sostanzialmente, dalle diverse sfumature di significato che possono non esistere tra le lingue (ad esempio nell'assegnazione del genere proprio della lingua italiana) oppure dai cambiamenti culturali avvenuti tra la pubblicazione del libro e la sua traduzione (come già menzionato, il testo è stato pubblicato nel 1991, quasi vent'anni prima della traduzione da noi effettuata).

Problematica	Soluzione
Traduzione di 'feminists'	Piuttosto che tradurre con 'femministe', si è preferito la personalizzazione del movimento, che non solo è una tecnica molto utilizzata nella lingua italiana, ma che permette di mettere completamente da parte il genere della parola.
Traduzione di 'gender confusion'	Abbiamo preferito il termine 'confusione dell'identità di genere', che rimanda a tutta una teoria assente nel testo di partenza per questioni cronologiche.
Presenza di termini generalmente considerati 'di cattivo gusto' (come parolacce)	Mantenuti in italiano

Tabella 2: *Risoluzione di problematiche di tipo etico*

Queste questioni sono state prese in considerazione nell'utilizzo dei software di traduzione automatica, e, per definizione, problematiche legate a questi ambiti sono state tutte considerate con una gravità molto bassa (in generale, 0), proprio perché non intaccano la comprensione del testo, ma sono 'preferenze' della cultura d'arrivo o del traduttore/cliente.

Sono state invece considerate problematiche con valore 1 quelle problematiche che intaccavano la lettura senza però complicare la comprensione del testo. Queste problematiche sono facilmente visibili e risolvibili dal lettore anche senza un'elevata conoscenza della lingua inglese o della disciplina. Queste problematiche sono quelle espressioni, ben note agli studenti di lingua straniera, che 'funzionano, ma non suonano bene'. Ne sono un esempio l'aggiunta di una *d* eufonica laddove l'italiano non ne sente il bisogno (*ad un altro*); l'uti-

lizzo di un apostrofo al posto dell'accento (*E' sempre sveglia*); mantenimento di un termine di lingua inglese laddove in italiano si può utilizzare sia questo che il termine italiano (*leadership*); utilizzo di una translitterazione non scientifica (*Mikhail Gorbaciov*); utilizzo di un articolo dal genere errato (*il jihad*⁹); mancanza della lettera maiuscola (*cristianesimo*); mancanza di spazi (*ea riscrivere*) e simili.

Valore 10 è stato assegnato a quelle problematiche che rendevano il testo difficile da leggere, senza però renderlo completamente incomprensibile. Problematiche di questo tipo possono spesso essere risolte, seppur con un certo lavoro da parte del lettore. Ne sono un esempio titoli di libri, di film o di canzoni lasciati in lingua originale (*The Territorial Imperative, Onward Christian Soldiers*); utilizzo di termini che, pur mantenendo il significato originale, non vengono utilizzati in italiano in quella specifica espressione (*e quindi "salta" in battaglia, a cercare i punti non custoditi*); espressioni letterali dall'inglese (*Dice poco, Il sollevarsi dell'oppresso*); traduzione di sinonimi con lo stesso termine, che porta a ripetizioni (*votare e votare*); utilizzo di una preposizione sbagliata (*stimola l'uomo (...) per compiere*); traduzione dello stesso termine inglese con termini diversi, laddove è utilizzato per motivi stilistici (*e l'esitazione (può portare) all'inerzia. L'inazione...*); il mancato utilizzo dell'aggettivo o avverbio 'proprio' (*porta ad un'intensa esperienza della sua vita*), e simili. Infine, il valore 100 è stato assegnato a problematiche che rendevano impossibile la comprensione del testo da parte del lettore, impossibili da risolvere senza la consultazione del testo originale e un certo livello di conoscenza della lingua inglese. Rientrano in questa categoria quindi, non solo passaggi che perdevano completamente significato, ma anche passaggi che assumevano significati assolutamente diversi da quello originale, fenomeno ancora più subdolo per il lettore, che non si accorgerebbe nemmeno della presenza di un problema. Sono esempi di problematiche di questo tipo traduzioni letterali di espressioni inglesi che rappresentano, come vedremo, la maggior parte delle problematiche riscontrate in corso d'opera (*soldati della fortuna* per *soldiers of fortune*, *shock da conchiglia* per *shell-shock*, *i tessuti sporchi* per *soiled tissues*); il mantenimento di forme inglesi non utilizzate in italiano (*Tyrant*); l'affibiazione di aggettivi al sostantivo sbagliato (*La storia della moglie del militare solitario e rifiutato*, laddove il testo originale definiva solitaria e rifiutata la moglie); inconsistenze nella traduzione (*..dei GI americani in Vietnam (...) in cui diversi IG ...*); disambiguazione errata di un termine con più traduzioni in lingua italiana (il termine *minister*, ad esempio, veniva sempre tradotto con *ministro* e mai con *sacerdote*, facendo pensare al ministro come figura politica piuttosto che come figura religiosa); mantenimento

⁹Bisogna notare che, in questo caso, almeno, la scelta del genere della parola è arbitraria. Se il termine, infatti, è di genere maschile in lingua araba, in italiano esso viene utilizzato molto più spesso al femminile.

di intere espressioni in lingua inglese (*"Forward, always forward, always forward"*); traduzione di nomi che non vengono tradotti in italiano (*permettendo a Saulo...*); mancata traduzione o traduzione errata di convenzioni del locale d'arrivo (*a.c.e* per *A.C.*¹⁰); traduzione di sezioni che dovevano rimanere in lingua inglese (*pronunciato "impasto"*, traduzione che non funziona perché il termine inglese *dough* ha ovviamente una pronuncia completamente diversa dalla proposta italiana, oppure *quello che il samurai chiamava fare*, dove il termine tradotto con *fare* era *do*, che, nonostante l'omografia con il verbo inglese, è un termine in lingua giapponese non segnalato); traduzione con una parte del discorso che fa perdere il significato all'espressione (*per poi ridiscendere - o decadimento*). Definite le problematiche secondo lo schema summenzionato, è possibile proseguire con l'analisi dei dati raccolti sui due traduttori automatici Google Translate e Deepl.

2.2 Analisi dei dati

Riprendiamo le formule proposte da dalla documentazione MQM per il calcolo della penalità del testo di arrivo:

$$P = (PM + PG * MG + PC * MC) / WC \quad (1)$$

e:

$$TQ = 100 - TP + SP$$

Ricordiamo inoltre che, per motivi di semplificazione, nel nostro caso la penalità per il testo originale (SP) avrà valore 0, motivo per cui possiamo semplificare la seconda formula in:

$$TQ = 100 - TP \quad (2)$$

La prima questione sarà dunque calcolare il *word-count* del testo tradotto. Nel caso l'utente usi un editor di testo come Word o OpenOfficeWriter, il conteggio delle parole sarà effettuato in automatico dal software, nel caso invece si usi (come nel nostro caso) \LaTeX , l'utente avrà diverse opzioni a sua disposizione per effettuare il conteggio:

- Utilizzo di un pacchetto come può essere TexCount;
- Utilizzo di uno strumento online, come ad esempio quello disponibile su wordcounter.io o wordcounter.net;
- Utilizzo di un codice JavaScript inserito in Adobe Reader.

¹⁰Interessante notare come la forma inglese b.c.e. sia probabilmente l'espressione che ha ricevuto più variabili di traduzioni (errate) tra i due traduttori automatici. Per citarne alcune: a.c.e.; a.c.e.e.; a.e.; bce; etc. Nel calcolo delle problematiche, ovviamente, non tutte queste varianti hanno avuto lo stesso valore: bce, pur essendo errata nell'uso italiano, è risolvibile dall'utente, seppur con un certo lavoro, e quindi vale come valore 10; a.c.e.e., d'altro canto, seppur risolvibile 'logicamente' quando si ritrova vicino una data, non è così semplice da comprendere per il lettore.

Nel nostro caso, il conteggio delle parole riscontrato nel testo originale e nelle traduzioni (effettuato, nel nostro caso, attraverso uno strumento online) è mostrato nella tabella seguente:

Sezione	Testo originale	Google Traduttore	DeepL
Totale	16267	15695	15764
Introduzione	1612	1564	1548
Il Re	3833	3727	3720
Il Guerriero	4910	4820	4849
Il Mago	3738	3561	3599
L'Amante	2174	2023	2048

Tabella 3: Conteggio delle parole del testo originale e della sua traduzione.

Nelle tabelle mostrate di seguito è possibile invece vedere il totale delle penalità, calcolate attraverso la formula 1, della sezione di testo tradotta e dei singoli capitoli, insieme al livello di qualità calcolato attraverso la formula 2.

Sezione	Google Traduttore	DeepL
Totale	2,23	2,01
Introduzione	1,85	1,02
Il Re	1,58	1,97
Il Guerriero	2,43	2,06
Il Mago	2,42	2,33
L'Amante	2,90	2,19

Tabella 4: Conteggio delle penalità.

Sezione	Google Traduttore	DeepL
Totale	97,77	97,99
Introduzione	98,15	98,98
Il Re	98,42	98,03
Il Guerriero	97,57	97,94
Il Mago	97,58	97,67
L'Amante	97,10	97,81

Tabella 5: Conteggio della qualità del testo.

Notiamo che, in generale, il software di traduzione DeepL ha ottenuto un punteggio di qualità maggiore rispetto al software di Google Traduttore, ottenendo un punteggio peggiore solo nel secondo capitolo. Ma quest'affermazione è troppo riduttiva per definire la qualità del lavoro dei due traduttori, e per accorgercene basta andare a calcolare il gli errori totali riscontrati e la media della loro gravità

Traduttore	Numero errori	Media della gravità
GT	589	59
DL	513	62

Tabella 6: Errori totali e media della gravità

Se è vero che ancora una volta il traduttore DeepL pare avere la meglio con un numero di errori inferiore a quelli riscontrati nella traduzione di Google Traduttore, è anche vero che la media ci dimostra che, in generale, questi errori hanno un peso maggiore sulla lettura del testo. Ma è anche vero che tale media superiore è dovuta al numero minore di errori piuttosto che al valore stesso degli errori, infatti gli errori riscontrati sono stati catalogati nella maniera seguente:

Traduttore	Gravità	Conteggio
GT	100	333
	10	158
	1	85
	0	13
DL	100	305
	10	123
	1	72
	0	13

Tabella 7: Conteggio degli errori per gravità.

In generale, si può dire che si sono riscontrati più errori nella traduzione di Google Traduttore rispetto a quella di DeepL. Questa differenza si nota, in particolare, per gli errori di gravità 100. In entrambi i casi, infatti, si tratta degli errori riscontrati con maggiore frequenza (più del doppio degli errori gravi per ambo i traduttori).

Nella figura 1 sono mostrate le tipologie di errori riscontrati tra i due traduttori, sulle quali è il caso di soffermarci per fare alcune osservazioni. Prima di tutto, si nota che alcune delle problematiche presenti nel nucleo MQM sono totalmente assenti dal testo tradotto. In generale, questo avviene poiché la preparazione del testo è avvenuta sotto lo sguardo di un traduttore umano: per questo motivo, problematiche come quelle di *design*, *internazionalizzazione* e *completezza* non potevano apparire. La terza problematica a non apparire è *legal-requirements* che, vista la tipologia testuale e l'assenza, almeno nelle sezioni tradotte, di specifiche legali, è necessariamente assente dalla nostra analisi. Se queste sono le problematiche completamente assenti, è anche interessante invece notare le problematiche che si sono presentate con maggior frequenza nel nostro lavoro. Uno sguardo ci basta a osservare che la tipologia *mistranslation* è quella maggiormente presente (tra i due traduttori si contano 513 problematiche di questo tipo, circa il 46% del totale, seguita da *fluency* (155 problematiche, circa il 14% del totale) e da *grammar* (123 problematiche di questo tipo, che ammontano all'11% del totale).

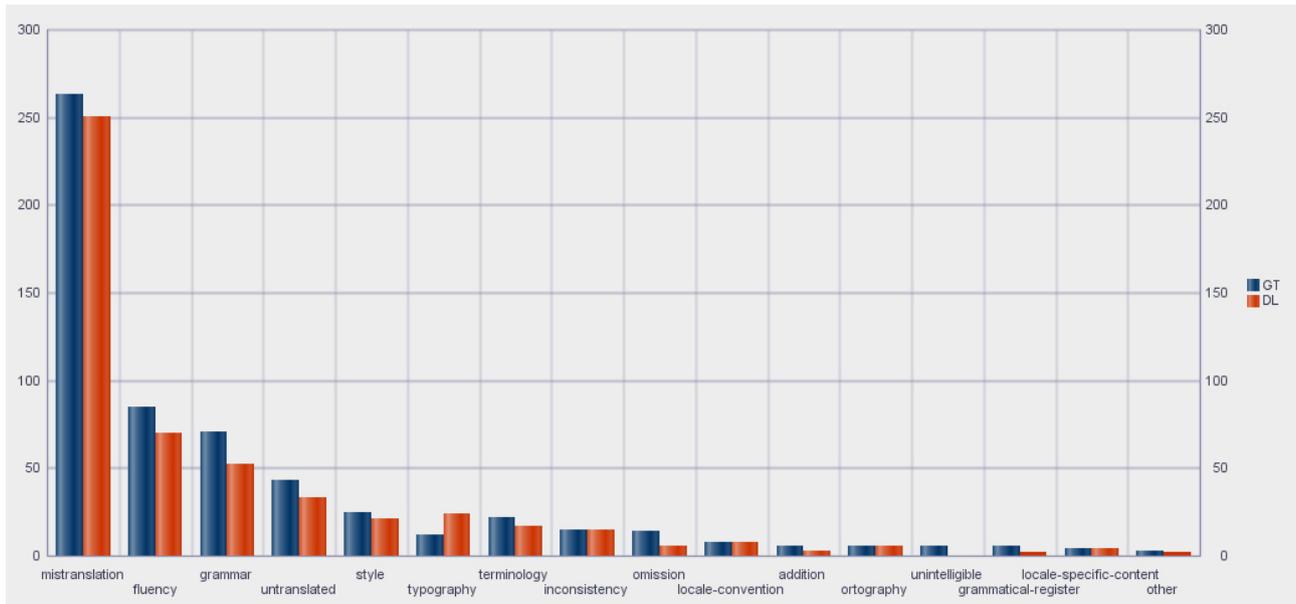


Figura 1: Tipologia degli errori tra Google Traduttore e DeepL

Ci soffermeremo quindi in particolare su queste tre tipologie nei paragrafi che seguono.

Mistranslation

Ricordiamo che l'analisi attraverso MQM può avere diverse tipologie di profondità, a seconda delle necessità del lavoro in corso. Nel nostro caso, ad esempio, è possibile andare più in profondità per quanto riguarda gli errori di tipo *mistranslation*, visto il peso che hanno avuto nella valutazione finale.

Un livello di profondità diverso indica, in sostanza, l'aggiunta di problematiche figlie, nel nostro caso:

- Traduzione ambigue (*ambiguous-translation*): la traduzione introduce un'ambiguità assente nel testo originale;
- Data/orario (*date-time*): non c'è concordanza tra le date o gli orari del testo originale e quelle del testo di arrivo;
- Entità, ovvero nomi propri o luoghi (*entity*): nomi propri o di luoghi non combaciano tra testo originale e testo tradotto;
- False friend (*false-friend*): nella traduzione appare erroneamente una parola solo superficialmente simile alla parola utilizzata nel testo originale;
- Traduzione errata di relazioni tecniche (*technical-relationship*): la traduzione, pur apparendo corretta dal punto di vista grammaticale, descrive una relazione tra oggetti che è errata nell'ambito spacialistico di cui si sta trattando;
- Numeri (*number*): i numeri non sono costanti tra testo originale e traduzione;
- Troppo letterale (*overly-literal*): il testo è stato tradotto in maniera molto letterale;
- Non dovrebbe essere stato tradotto (*no-translate*): sezioni che sarebbero dovute rimanere in lingua originale sono state tradotte;

- Conversione di unità di misura (*unit-conversion*): il testo tradotto non presenta le conversioni necessarie delle unità di misura che invece ci si aspetterebbe.

Possiamo dire che, in generale, la tipologia di *mistranslation* maggiormente presente nelle nostre analisi è *overly-literal*. È indubbio che il traduttore automatico, essendo una macchina, non ha la possibilità (ancora, perlomeno) di intendere in quale situazione un termine viene utilizzato con un suo significato traslato e tenderà, generalmente, a tradurlo con il suo significato diretto. Questo è vero per molte parole lessicali, come ad esempio:

- Draw his sword : disegna la sua spada invece di *estrae la sua spada*;
- The planes : gli aerei, laddove il testo faceva invece riferimento ai *piani* della realtà;
- Indiana Jones is racing the Nazis : sta correndo i nazisti;
- is going : è partita invece di *è azionata*;
- the witch burnings : bruciature delle streghe, invece di *roggi*;
- "straight" professions : professioni "dirette", invece di *professioni "normali"*;
- blues, oranges, whites, reds, and maroons : blu, arance, bianchi, rossi, in cui il traduttore non è riuscito ad evincere dal contesto che, nonostante il plurale, si trattava di un colore e non del frutto;
- body locks : serrature del corpo, invece di *prese*;
- interview : intervista, invece di *colloquio*.

Così come per alcune intere espressioni idiomatiche, che vanno a perdere completamente il loro significato:

- Save the day : salva la giornata, invece di *risolve la situazione*;

- in terms little short of libidinous : in termini poco brevi di libido, invece di *termini a dir poco libidinosi*;
- the story goes : così va la storia, e non *come vuole la leggenda*;
- soldiers of fortune : soldati della fortuna, e non *mercanari*;
- treading the water : calpestare l'acqua, invece di *tentare di restare a galla*;
- broke new ground : ha rotto un nuovo terreno, invece di *aprire nuove strade*.

e per alcuni false friend come:

- Feast: festa;
- Minister: ministro, invece di sacerdote (seppur l'utilizzo del termine 'ministro' per indicare un uomo di Chiesa sia utilizzabile nella lingua italiana, esso è spesso accompagnato da 'di Chiesa', e il suo utilizzo senza specificazioni riporta, più comunemente, alla sua valenza politica).

Ma se questi errori possono essere comprensibili, in parte soprattutto perché appaiono anche nelle prime fasi di apprendimento di una lingua straniera, non è stato sempre possibile individuare chiaramente la causa di una traduzione errata:

- Canaanite god Baal's love : l'amore del dio canadese, invece di *cananeo*;
- Christ child : il figlio di Cristo;
- The weak father : il debole del padre;
- show the sad fact : mostrano il dispiacere;
- ne'er-do-well dissidents : dissidenti benestanti.

Alcuni di questi errori possono essere riportati ai già menzionati problemi relativi alla risoluzione, da parte dei software ad architettura NMT, di problematiche relative alla traduzione di termini nuovi o rari.

Fluency

Le problematiche di questo tipo sono le seconde per frequenza. In questo caso, il nucleo MQM offriva già delle problematiche più specifiche riguardo la fruibilità del testo, ma nella nostra analisi la problematica genitore ha riscontrato una frequenza molto più elevata di quelle figlie. Questa tipologia, come viene descritto dalla documentazione ufficiale MQM, è relativa alle problematiche possono essere identificate anche senza effettuare comparazioni tra il testo originale e il testo tradotto, essendo indipendenti dallo status di traduzione del testo analizzato.

Molte di queste problematiche riguardavano la scelta di termini che creavano un certo rumore nella traduzione, come ad esempio:

- praticare la sopravvivenza;
- con una mano di ferro;
- un nemico travolgente;
- impartiva loro segreti.

In generale, questo tipo di problematiche raramente ha raggiunto una gravità 100 (42 esempi su un totale di 155), e in generale definiva sezioni di testo che potevano 'suonare male' nella lettura del testo, ma senza intaccarne completamente la comprensione.

Grammar

Le problematiche riscontrate in questa categoria possono essere riportate a tre sottotipologie differenti, classificabili nelle tre sottotipologie:

- function-words (ovvero uso incorretto, aggiunta o omissione nella traduzione di una parola funzionale);
- word-form (ovvero un problema con la forma grammaticale di una parola);
- word-order (ovvero relativo all'ordine delle parole).

Problemi relativi alle parole funzionali sono stati ad esempio:

- (...) exploits and abuses others : sfrutta e abusa degli altri;
- (...) starved for King energy : affamato per l'energia del Re;
- It is to him that people go (...) : È a lui che le persone vanno.

In generale, possiamo dire che in questa sottocategoria rientrano la maggior parte delle problematiche di tipo grammaticale.

Un discorso a parte andrebbe fatto per la traduzione dei possessivi, che, anche nel caso di identità soggetto e possessore dell'oggetto, non sono stati praticamente mai tradotti con l'aggettivo *proprio*.

Per quanto riguarda la forma grammaticale delle parole invece:

- Not only does the fencer train his body, (...) : Non solo lo schermidore addestrano il suo corpo;
- It was the magicians (...) who created civilization (...) : Furono i maghi (...) che creò la civiltà;
- and touches the Ark to steady it.: e tocca l'Arca per fermarlo.

In questa categoria sono inoltre rientrati tutti gli errori relativi all'aspetto verbale:

- Native American men lived and died with the Warrior energy informing even the smallest of their acts (...) : i nativi americani vissero e morirono con l'energia del Guerriero informando anche i più piccoli dei loro atti;
- And it also proclaimed that there was only one direction : E ha anche proclamato che c'era una sola direzione(...).

Alcuni problemi, ovviamente, sono qui derivati dalla presenza nella lingua italiana di forme grammaticali assenti invece nella lingua inglese, come le marcature di genere. Altri problemi sono invece sorti in situazioni

difficilmente spiegabili, come soggetti plurali accompagnati da verbi al singolare (anche se, in alcuni casi, questo è stato causato dalla distanza che c'era tra soggetto e verbo, come possiamo vedere nel secondo esempio tra quelli sopra citati).

Infine, problematiche grammaticali sono comparse in relazione all'ordine delle parole nella traduzione:

- per miracolosamente trasformare;
- un gentile zio.

In ogni caso, a prescindere dalla gravità di queste problematiche, si sono rivelate, insieme a quelle della tipologia più generale *fluency*, le più semplici da individuare.

Conclusioni

Lo standard MQM si è rivelato uno degli strumenti più utili nel corso del lavoro svolto, in particolare per le possibilità che offre nei confronti di un eventuale post-editing successivo, per il quale permette di avere una classificazione e categorizzazione degli errori di enorme valore. Nel mondo attuale, in cui i dati e la loro analisi sono diventati essenziali per qualunque professione, un'analisi di questo tipo, standardizzata e riutilizzabile, è di valore innegabile.

Bibliografia

- Arle, Lommel et al. (2015). *Multidimensional Quality Metrics (MQM) Issue Types*. URL: {<http://www.qt21.eu/mqm-definition/issues-list-2015-12-30.html>}. {Consultato il 11/11/2018}.
- Bray, Tim et al. (2009). *Namespaces in XML 1.0 (Third Edition)*. URL: {<https://www.w3.org/TR/REC-xml-names/#URIRef>}. {Consultato il 07/11/2018}.
- Harris, Kim et al. (2015). *Multidimensional Quality Metrics (MQM) Definition*. A cura di Arle Lommel (DFKI), Aljoscha Burchardt (DFKI), Hans Uszko-reit (DFKI). URL: {<http://www.qt21.eu/mqm-definition/definition-2015-12-30.html>}. {Consultato il 23/10/2018}.
- Vermeer, Hans Josef e Katharina Reiß (1984). *Grundlegung einer allgemeinen Translationstheorie*. A cura di Tübingen, Niemeye.