

Lara Bernardi

Il pre-editing e la traduzione automatica

Un linguaggio controllato per scrivere testi in italiano
adatti alla traduzione automatica

Ricerca sulla traduzione automatica

Lo scopo della mia ricerca è stato quello di trovare delle regole da applicare ai testi in lingua italiana per renderli adatti alla traduzione automatica. Ho analizzato testi tecnico-descrittivi sul tema delle energie rinnovabili e le relative traduzioni in tedesco, francese e inglese. Ho svolto un lavoro di pre-editing sui testi originali (sulla base delle regole di Muegge per un linguaggio controllato e sugli studi di Nielsen per una scrittura adatta al web) e ho evidenziato gli aspetti principali da tenere in considerazione quando si intende sottoporre un testo scritto in italiano a un sistema di traduzione automatica. È necessario considerare sia aspetti puramente linguistici sia aspetti legati al significato e alla struttura delle frasi. Di seguito elenco le regole principali da osservare.

1. Evitare errori ortografici e di punteggiatura (in particolare apostrofi e accenti)

Le macchine non pensano come gli uomini: se noi leggiamo “conquistare” possiamo dedurre senza particolari problemi che è presente un errore e che il verbo corretto è “conquistare”. Il sistema di traduzione automatica non ragiona in questo modo e non proporrà nessuna traduzione soddisfacente. La questione della punteggiatura, in particolare quella della virgola (dove si trova, davanti o dopo quale costruzione sintattica) è una questione ancora controversa e non soggetta a regole, se non quella di evitarne un uso troppo frequente.

2. Scrivere frasi con una struttura semplice, con il soggetto al primo posto e il verbo al secondo posto

Dalle mie analisi è emerso che qualsiasi tipo di frase con qualsiasi tipo di contenuto viene tradotta meglio se presenta una struttura con il soggetto al primo posto e il verbo in seconda posizione. Il soggetto e il verbo sono parti fondamentali di una frase e nella maggior parte dei casi, se presenti, vengono tradotti e posizionati in modo corretto.

3. Evitare frasi troppo lunghe che contengano più di due soggetti o più di due verbi

Le frasi troppo lunghe rappresentano un grande problema per i sistemi di traduzione automatica poiché, nella maggior parte dei casi, sono composte da coordinate e subordinate che presentano più soggetti e più verbi o verbi diversi che dipendono dallo stesso soggetto. Già un traduttore esperto potrebbe trovarsi in difficoltà nel riportare una frase lunga in un'altra lingua, ma la rilettura e la scomposizione mentale della frase permettono al traduttore di trovare una soluzione ideale. Il computer non svolge però nessun lavoro di scomposizione della frase. L'elenco puntato e la scomposizione della frase lunga in frasi semplici più corte possono aiutare.

4. Preferire frasi coordinate a frasi subordinate

In seguito alle mie ricerche sono giunta alla conclusione che le frasi adatte alla traduzione automatica sono frasi semplici al presente e al passato prossimo (eventualmente anche al futuro se è necessario) con il soggetto in prima posizione e il verbo in seconda posizione. Dato che è poco probabile riuscire a scrivere un intero testo composto unicamente da frasi semplici, per quanto riguarda le frasi coordinate i tempi verbali più adatti da utilizzare sono sempre il presente e il passato prossimo. È necessario ripetere l'intera forma verbale in tutte le frasi e non inserire troppe informazioni tra un verbo e l'altro se essi dipendono dallo stesso soggetto citato una volta sola. Tra le frasi subordinate, le finali e le concessive risultano migliori rispetto alle temporali e alle causali. Ovviamente, più le frasi sono corte, meglio verranno tradotte. Questo comporta però uno svantaggio a livello di corposità del testo: un testo composto da frasi corte e semplici non avrà lo stesso effetto di un testo più articolato e che presenta strutture subordinate. Tuttavia, se lo scopo finale è la traduzione automatica, vale la pena penalizzare un po' la corposità del testo di partenza in favore di una sua più accurata traduzione.

5. Inserire sempre gli articoli davanti ai sostantivi, in modo tale che i sostantivi siano riconoscibili

L'inserimento degli articoli davanti ai sostantivi è fondamentale in modo tale che il sistema di traduzione automatica non confonda il sostantivo con un'altra parte della frase, come successo ad esempio nella traduzione della frase seguente.

	<i>Testo originale</i>	<i>Testo modificato il 02.04.2015</i>	<i>Testo modificato il 02.04.2015</i>
<i>Italiano</i>	[...] consentendone un maggior utilizzo durante l'anno grazie all'energia solare. ¹	Utilizzo annuale maggiore grazie all'energia solare	L'energia solare permette un utilizzo annuale maggiore
<i>Francese</i>	[...] permettant une plus grande utilisation toute l'année grâce à l'énergie solaire.	Utilisez les accrue grâce à l'énergie solaire annuels	L'énergie solaire permet une utilisation annuelle supérieure
<i>Tedesco</i>	[...] die in verstärktem Maße Jahres dank Solarenergie überall.	Verwenden Sie die jährlichen erhöht dank Solarenergie	Solarenergie ermöglicht eine höhere jährliche Nutzung
<i>Inglese</i>	[...] allowing greater use throughout the year thanks to solar energy.	Use the annual increased thanks to solar energy	Solar power allows a higher annual use

¹ Per questioni di spazio, all'interno delle tabelle ho eliminato alcune parti delle frasi non necessarie per comprendere i miei ragionamenti e le mie modifiche. In Google translate è però stata inserita la frase per intero.

6. Evitare i pronomi e inserire sempre il soggetto della frase, se possibile al primo posto

L'uso dei pronomi è altamente sconsigliato: per quanto possa essere utile in un testo originale per evitare la ripetizione di un nome magari appena citato nella frase precedente, nella fase di traduzione automatica il pronome viene considerato esattamente come una particella traducibile, ma che non riprende nessun tipo di significato e che nella maggior parte dei casi fa risultare incorretta la frase. È quindi molto importante inserire il soggetto, se possibile all'inizio della frase.

	<i>Testo originale</i>	<i>Testo modificato 02.04.2015</i>
<i>Italiano</i>	Particolarmente adatti per riscaldare piscine scoperte ad uso estivo, aumentano di alcuni gradi la temperatura dell'acqua, [...]	I collettori solari sono adatti per il riscaldamento di piscine scoperte utilizzate in estate: • La temperatura dell'acqua aumenta di alcuni gradi [...]
<i>Francese</i>	Particulièrement adapté pour le chauffage des piscines extérieures pour l'été, ont augmenté de degrés la température de l'eau quelques-uns, [...]	Les capteurs solaires sont adaptés pour le chauffage des piscines extérieures utilisées en été: • La température de l'eau augmente de quelques degrés [...]
<i>Tedesco</i>	Besonders geeignet für die Heizung Außenpools für den Sommer, erhöht um ein paar Grad Wassertemperatur, [...]	Die Solarkollektoren sind für die Beheizung von Außenpools im Sommer genutzt: • Die Wassertemperatur steigt ein paar Grad [...]
<i>Inglese</i>	Particularly suitable for heating outdoor swimming pools for summer use, increased by a few degrees water temperature, [...]	The solar collectors are suitable for the heating of outdoor pools used in summer: • The water temperature increases a few degrees [...]

7. Collegare sempre, e soprattutto visibilmente, elementi che dipendono tra loro (ad es. *preposizione + sostantivo, verbo + preposizione, sostantivo + aggettivo ecc.*), evitando così collegamenti logici solo tramite un ragionamento umano

Se più parole o parti della frase dipendono dalla stessa preposizione, è necessario collegare le varie parti in modo “visibile” e non solo in modo logico. È ad esempio utile ripetere la preposizione davanti ad ogni parola o parte che dipende da essa. Il sistema di traduzione automatica non è in grado di comprendere il collegamento logico, ovvio per una persona, tra una preposizione e le parole che dipendono da quest'ultima.

Nel mio corpus era ad esempio presente una frase che conteneva “tubo di gomma, polietilene o metallico”. In primo luogo in questo caso va corretto “metallico” con “metallo” se vogliamo legarlo a “tubo di” e in secondo luogo è necessario inserire la preposizione “di” davanti a “polietilene” e a “metallo”. Solo in questo modo il sistema di traduzione automatica sarà in grado di riportare nel migliore dei modi il significato in un'altra lingua.

8. Se una preposizione dipende da un verbo presente nella frase, è necessario inserirla subito dopo il verbo

Nella fase di pre-editing ho notato che se all'interno della frase la preposizione retta dal verbo viene inserita (subito) dopo il verbo, la traduzione risultava maggiormente corretta. Questo e altri tipi di legame devono essere evidenti e il più semplice possibile nella frase.

Testo 2.1	Testo originale (r. 1)	Testo modificato il 14.04.2015
Italiano	DA UN'INTUIZIONE NASCE SOLARTEC	Solartec nasce da un'intuizione
Francese	Intuition BORN SOLARTEC	Solartec vient de l'intuition
Tedesco	Intuition BORN SOLARTEC	Solartec stammt aus einer Intuition
Inglese	Intuition BORN SOLARTEC	Solartec comes from an intuition

9. Se più aggettivi definiscono un sostantivo, è necessario ripetere il sostantivo davanti a ogni aggettivo

È fondamentale legare ogni singolo aggettivo al sostantivo che esso definisce. Un elenco di più aggettivi separati da virgole, che dipende da un unico sostantivo, non sarà riconosciuto tale da Google translate. La possibilità che il sistema riesca a tradurre gli aggettivi in modo corretto è più elevata se si traduce in francese poiché la struttura italiana *sostantivo + aggettivo*, + *aggettivo* rimane tale. In tedesco e in inglese, al contrario, gli aggettivi vengono sempre scritti prima del sostantivo e quindi vi è un maggior rischio che il primo aggettivo venga posto correttamente davanti al sostantivo, ma che gli altri vengano tradotti come aggettivi a sé stanti o come sostantivi.

	Testo originale	Testo modificato il 07.04.2015
Italiano	Tutti i moduli sono sottoposti ad un controllo visivo, meccanico ed elettrico.	Tutti i moduli sono sottoposti a vari controlli: <ul style="list-style-type: none"> • un controllo visivo • un controllo meccanico • un controllo elettrico.
Francese	Tous les modules sont soumis à une inspection visuelle, mécanique et électrique.	Tous les modules sont soumis à différents contrôles: <ul style="list-style-type: none"> • une inspection visuelle • une commande mécanique • un contrôle électrique.
Tedesco	Alle Module sind mit einer visuellen Inspektion, mechanische und elektrische unterworfen.	Alle Module werden verschiedenen Prüfungen unterzogen: <ul style="list-style-type: none"> • eine visuelle Inspektion • eine mechanische Steuerung • eine elektrische Prüfung.
Inglese	All modules are subjected to a visual inspection, mechanical and electrical.	All modules are subjected to various checks: <ul style="list-style-type: none"> • a visual inspection • a mechanical control • an electrical check.

10. Evitare i sinonimi e non temere le ripetizioni: usare le stesse parole per esprimere lo stesso concetto

Il perché di questa regola è molto semplice da comprendere: già Muegge era giunto a tale conclusione. La ricerca di parole conosciute e presenti nel linguaggio comune permette di aumentare la probabilità che il sistema di traduzione automatica traduca correttamente una determinata parola. È sbagliato pensare che due sinonimi in una lingua di partenza vengano tradotti con un unico termine nella lingua di arrivo, come magari farebbe un traduttore umano per rendere il testo maggiormente comprensibile e privo di incertezze. Essendo una macchina, è molto probabile che Google translate traduca parole diverse in modo diverso, creando così confusione. La ripetizione renderà il testo poco originale, ma al contempo la traduzione maggiormente precisa e il processo di traduzione semplice.

11. Evitare il gerundio, le forme passive, i verbi riflessivi e i verbi impersonali

Il gerundio è una forma verbale molto particolare: non esiste in tutte le lingue ed è quindi necessario evitare il suo uso. Vi sono vari modi per sostituire il gerundio.

	<i>Testo originale</i>	<i>Testo modificato il 02.04.2015</i>
<i>Italiano</i>	“I collettori solari per piscine TS POOL, lavorando a bassa temperatura, sono [...]”	I collettori solari sono [...], perché lavorano a basse temperature.
<i>Francese</i>	Les capteurs solaires pour piscines TS POOL, travaillant à basse température, [...]	Les capteurs solaires sont [...], parce qu'ils travaillent à des températures basses.
<i>Tedesco</i>	Die Solarkollektoren für Schwimmbäder TS POOL, arbeitet bei niedrigen Temperaturen, [...]	Sonnenkollektoren sind [...], weil sie bei niedrigen Temperaturen zu arbeiten.
<i>Inglese</i>	The solar collectors for swimming pools TS POOL, working at low temperature, [...]	Solar collectors are [...], because they work at low temperatures.

I verbi al gerundio vengono spesso tradotti in modo corretto solo in inglese, poiché corrispondono alle forme –ing, usate molto frequentemente.

La problematica delle frasi impersonali è invece legata a quella del soggetto: il sistema di traduzione automatica non traduce nessuna frase senza collegare il verbo a un soggetto e, se nel testo originale il soggetto non c'è, il sistema di traduzione proporrà una soluzione sicuramente sbagliata legando il verbo a un altro sostantivo o aggiungendo un pronome casuale. È quindi necessario evitare le forme impersonali e inserire sempre il soggetto all'interno della frase (ancora meglio se al primo posto). Nella frase “[...] per ogni kWh prodotto si risparmiano circa 250gr. di olio combustibile.” si riscontra il problema della forma impersonale e nelle traduzioni Google translate propone dei verbi principalmente al futuro: in francese “[...] pour chaque kWh produit

permettra d'économiser environ 250g. de fuel-oil.”, in tedesco “[...] für jede kWh produziert wird ca. 250g speichern. Heizöl.” e in inglese “[...] for each kWh produced will save approximately 250g. of fuel oil.”. Nella fase di pre-editing ho inserito il soggetto “sistema” e ho mantenuto il contenuto originale della frase: “Il sistema risparmia 250 grammi di olio combustibile per ogni kWh.”.

12. Favorire elenchi a frasi lunghe che presentano vari elementi suddivisi da virgole

L'elenco puntato è uno degli strumenti introdotti da Nielsen nelle sue ricerche sulla scrittura per il web che ho sfruttato maggiormente nella fase di pre-editing per rendere più semplici frasi lunghe e ricche di virgole presenti nei miei testi. L'uso delle virgole non è infatti uguale in tutte le lingue; l'elenco puntato permette invece di trasformare una frase lunga in frasi più corte o in frasi nominali.

	<i>Testo originale</i>	<i>Testo modificato il 07.04.2015</i>
<i>Italiano</i>	Moduli fotovoltaici policristallini con alte prestazioni, celle di silicio solare (156.5 mm ²), efficienza del modulo fino al 15,2 %.	<ul style="list-style-type: none"> • moduli fotovoltaici policristallini con alte prestazioni • celle di silicio solare (156.5 mm²) • efficienza del modulo fino a 15,2 %.
<i>Francese</i>	Modules photovoltaïques polycristallins avec des cellules de silicium de haute performance solaires (156,5 mm ²), module de rendement allant jusqu'à 15,2%.	<ul style="list-style-type: none"> • Modules photovoltaïques polycristallins à haute performance • cellules solaires en silicium (156,5 mm²) • Module efficacité jusqu'à 15,2%.
<i>Tedesco</i>	Polykristalline Photovoltaik-Module mit Hochleistungs-Silizium-Solarzellen (156,5 mm ²), Modulwirkungsgrad von bis zu 15,2%.	<ul style="list-style-type: none"> • Polykristalline Photovoltaik-Module mit hoher Leistung • Silizium-Solarzellen (156,5 mm²) • Modulwirkungsgrad bis zu 15,2%.
<i>Inglese</i>	Polycrystalline photovoltaic modules with high-performance silicon solar cells (156.5 mm ²), module efficiency of up to 15.2%.	<ul style="list-style-type: none"> • Polycrystalline photovoltaic modules with high performance • silicon solar cells (156.5 mm²) • module efficiency up to 15.2%.

13. Evitare di inserire locuzioni temporali al primo posto nella frase

Sempre per permettere al soggetto di essere posizionato al primo posto nella frase, consiglio di posizionare le locuzioni temporali in mezzo o alla fine di una frase.

	<i>Testo modificato il 07.05.2015</i>	<i>Testo modificato il 07.05.2015</i>	<i>Testo modificato il 07.05.2015</i>	<i>Testo modificato il 07.05.2015</i>
<i>Italiano</i>	Negli ultimi anni Solterra realizza impianti fotovoltaici.	Solterra realizza impianti fotovoltaici dal 2010 .	Solterra realizza impianti fotovoltaici da diversi anni .	Solterra realizza impianti fotovoltaici a partire dal 2010 .
<i>Francese</i>	Au cours des dernières années Solterra fabrique des systèmes photovoltaïques.	Solterra fabrique des systèmes photovoltaïques depuis 2010.	Solterra fabrique des systèmes photovoltaïques depuis plusieurs années.	Solterra fabrique des systèmes photovoltaïques depuis 2010.
<i>Tedesco</i>	In den letzten Jahren Solterra	Solterra fertigt Photovoltaikanlag	Solterra fertigt Photovoltaikanlag	Solterra fertigt Photovoltaikanlage

	fertigt Photovoltaikanlagen.	en seit 2010.	en für mehrere Jahre.	n seit 2010.
<i>Inglese</i>	In recent years Solterra manufactures photovoltaic systems.	Solterra manufactures photovoltaic systems since 2010.	Solterra manufactures photovoltaic systems for several years.	Solterra manufactures photovoltaic systems since 2010.

14. Evitare una terminologia troppo ricercata per esprimere un concetto che può essere spiegato con parole più semplici ed evitare formulazioni letterarie o a effetto che però non riportano nessun significato rilevante

La frase “Albasolar affonda le sue radici nel 1968, [...]” presente nel mio corpus è un tipico esempio di ciò che non va proposto a un sistema di traduzione automatica. Tutto ciò che crea colore e ricchezza lessicale va portato a un livello inferiore e sostituito da una terminologia più comune. In questo caso la frase potrebbe essere semplificata in ampio modo per risultare maggiormente adatta alla traduzione automatica, per esempio così: “Albasolar nasce nel 1968, [...]”.

15. Evitare qualsiasi parola straniera

Per ovvi motivi, se possibile, le parole straniere vanno tralasciate. Probabilmente in futuro, tramite l’inserimento sempre più elevato di testi di vario genere, Google translate sarà in grado di comprendere che parole come ad esempio “governance” sono ormai diventate di uso comune anche in altre lingue, ma per ora consiglio di tralasciarle e di trovare corrispettivi nella lingua originale, per evitare problemi supplementari.

16. Evitare il *non* espletivo

Il *non* espletivo o pleonastico rappresenta la presenza di un *non* in una frase senza che ne alteri il significato. Il *non* è infatti facoltativo e ha una funzione non del tutto chiara, a volte fa risaltare la frase. Si ritrova molto spesso in proposizioni temporali introdotte da “finché”. Nelle altre lingue questo *non* viene però tradotto come una negazione. È quindi importante evitare di scrivere frasi di questo tipo.

	<i>Frase</i>
<i>Italiano</i>	La produzione avverrà in Germania finché gli stabili in Svizzera non saranno pronti .
<i>Francese</i>	La production aura lieu en Allemagne jusqu'à ce que les bâtiments en Suisse ne sera pas prêt .
<i>Tedesco</i>	Die Produktion wird in Deutschland bis in die Gebäude in der Schweiz stattfinden wird nicht bereit sein .
<i>Inglese</i>	The production will take place in Germany until the buildings in Switzerland will not be ready .