



FRANCESCO MELEDANDRI

**COMPRENDERE I SISTEMI
DI TRADUZIONE AUTOMATICA
ANALISI QUALITATIVE SU TESTI IT><EN**





©

ISBN
979-12-218-0504-8

PRIMA EDIZIONE
ROMA 31 GENNAIO 2023

INDICE

- 7 *Presentazione del lavoro*
- 13 *Introduzione*
1. Il rapporto fra uomo e macchina: evoluzione cronologica, 13 – 11. Una macchina in grado di tradurre: la *Machine Translation*, 19 – 111. Problematiche di resa traduttiva: la distanza interlinguistica, 22.
- 29 Capitolo I
Storia ed evoluzione dei primi approcci alla *Machine Translation*
1.1. Fasi storiche della *Machine Translation*, 29 – 1.2. La fase embrionale: macchine e calcolatori, 31 – 1.3. Il sogno (semi)meccanico: le macchine di Trojanskij e di Artsrouni, 36 – 1.4. Enigma: una “macchina per tradurre”, 42 – 1.5. Il dopoguerra e la nascita della *Machine Translation* su basi scientifiche, 48 – 1.6. Il contributo di Warren Weaver e l’importanza del suo “memorandum”, 53 – 1.7. La scuola americana: il MIT e il ruolo dei “pionieri” della *Machine Translation*, 55 – 1.8. Bar-Hillel e la definizione degli obiettivi per la MT, 58 – 1.9. Altre forme di traduzione: la MT “impura”, 63 – 1.10. La Grammatica Universale, 68 – 1.11. La dimostrazione della “non-feasibility” della “pure MT”, 70 – 1.12. Il Rapporto *Alpac*: la fine del sogno meccanico, 73.
- 81 Capitolo II
Caratteristiche peculiari della *Machine Translation* e nuovi approcci traduttivi
2.1. Il rapporto di mimesi tra uomo e macchina, 81 – 2.2. Schemi “classici”, nuovi punti di vista, 86 – 2.3. Gli esordi: il metodo diretto, 92 – 2.4. Verso

un sistema universale: l'approccio *rule-based* e le metodologie *transfer-based* e *interlingua-based*, 97 – 2.5. La rivoluzione dei *Big Data*: l'approccio statistico dei sistemi di *Machine Translation*, 107 – 2.6. Verso un sistema “meccanicamente intelligente”: la Traduzione Automatica Neurale, 113.

121 *Capitolo III*

Valutazione comparativa di testi IT><EN attraverso un *assessment framework*

3.1. Traduzioni umane e artificiali: nuovi ruoli, nuovi attanti, 121 – 3.2. L'importanza di Internet nei sistemi di Traduzione Automatica, 126 – 3.3. I sistemi di Traduzione Automatica presenti sul Web: una risorsa di comunicazione globale, 128 – 3.4. I sistemi di *Web-Based Machine Translation* commerciali e gratuiti oggetto di analisi, 135 – 3.5. Procedura di valutazione di testi sottoposti a Traduzione Automatica, 137 – 3.6. Metodologia di analisi, 144 – 3.7. Tipologie di testi sottoposti a Traduzione Automatica, 146 – 3.8. Criterio di valutazione delle Traduzioni Automatiche analizzate, 152 – 3.9. Risultati, 156 – 3.9.1. *Coppia Linguistica: Italiano–Inglese*, 156 – 3.9.1.1. *Testo 1: Ricorso legale*, 156 – 3.9.1.2. *Testo 2: Manuale di istruzioni*, 157 – 3.9.1.3. *Testo 3: Estratto di Annual Report*, 158 – 3.9.2. *Coppia Linguistica: Inglese–Italiano*, 159 – 3.9.2.1. *Testo 1: Ricorso legale*, 159 – 3.9.2.2. *Testo 2: Manuale di istruzioni*, 160 – 3.9.3. *Testo 3: Estratto di Annual Report*, 161 – 3.9.3. *Tabelle di confronto relative ai singoli sistemi analizzati*, 162 – 3.9.4. *Tabella riassuntiva di valutazione in seguito all'applicazione di modificatore di valutazione*, 165 – 3.10. Considerazioni finali: qualità della Traduzione Automatica *Web-Based* e gratuita, 166

177 *Bibliografia*201 *Appendice*

Testi sottoposti a procedura di *Machine Translation*

1. Testi di partenza – Inglese, 199 – 1.1. *Testo 1: Ricorso Legale*, 199 – 1.2. *Testo 2: Manuale di istruzioni*, 201 – 1.3. *Testo 3: Estratto di Annual Report*, 202 – 2. Testi di partenza – Italiano, 206 – 2.1. *Testo 1: Ricorso Legale*, 203 – 2.2. *Testo 2: Manuale di istruzioni*, 205 – 2.3. *Testo 3: Estratto di Annual Report*, 206.

PRESENTAZIONE DEL LAVORO

Il presente lavoro intende analizzare alcune problematiche specifiche nel campo della traduzione, con particolare riferimento alla traduzione tecnico-specialistica. L'attenzione sarà rivolta all'ambito relativo alla traduzione automatica di testi e, in senso più ampio, al modo di poter comprendere tali sistemi per un utilizzo consapevole e in vari contesti d'uso da parte di traduttori e non.

Il Capitolo I del lavoro sarà dedicato all'evoluzione della Traduzione Automatica nel corso degli ultimi decenni: si procederà ad una sorta di cronologia degli eventi, cercando di sottolineare l'importanza delle scoperte tecnologiche (e delle relative migliorie in ambito linguistico) nel campo dell'*Information Technology* che hanno portato di fatto alla crescita in termini di resa linguistica delle traduzioni *non-natural*. Si delinearanno anche le motivazioni di fondo che hanno spinto vari team scientifici alla creazione di interfacce sempre più evolute, veloci e precise. Particolare attenzione sarà dedicata anche al noto report del 1966 stilato dall'*Automatic Language Processing Advisory Committee*⁽¹⁾ nel quale si delinea, in maniera piuttosto definitiva per quel tempo (ma il documento, alla luce dei risultati più moderni, sembra più che mai attuale), lo stallo della situazione della qualità di output dei testi tradotti in automatico.

(1) *Language and machines: computers in translation and linguistics. A report by the Automatic Language Processing Advisory Committee, Division of Behavioral Sciences, National Academy of Sciences, National Research Council*, Publication 1416, National Academy of Sciences, National Research Council 1966, Washington D.C. 1966.

Nel Capitolo II si passeranno in rassegna i nuovi approcci linguistici applicati alle macchine: dal sistema di pura sostituzione, che rappresenta la base dell'approccio *Machine Translation* ma anche un sistema "chiuso", si è passati — in special modo nel corso degli ultimi venti anni — a sistemi più aperti e meno vincolati al *word level*. I metodi *example-based* e *corpus-based* sono le nuove basi su cui si poggiano i nuovi studi nel campo della Traduzione Automatica, poiché questi, come accennato in precedenza, espandono il criterio di individuazione della *source language* e della resa traduttiva nella *target language* a un livello superiore a quello precedente. Si passa dal *word level*, dunque un approccio prettamente logocentrico, ad un orientamento *sentence level*, con un focus incentrato sulla componente sintattica dei testi, che a sua volta non può non considerare anche la componente lessicale. Il *text level*, pur essendo di difficile implementazione a causa delle numerose variabili in gioco nel processo traduttivo⁽²⁾ rappresenta la nuova frontiera dei sistemi di Traduzione Automatica, in quanto i nuovi sistemi di sintesi basati sulle cosiddette Reti Neurali ha spostato il focus principale verso un vero e proprio riconoscimento globale dei con-testi da parte delle macchine, avvicinandosi sempre più agli stessi schemi logico-sequenziali e di associazioni ontologiche di idee del cervello umano.

Focalizzando l'attenzione sulla qualità traduttiva di output, il Capitolo III del presente lavoro intende effettuare un'analisi linguistica di testi tradotti artificialmente. In altre parole, verrà effettuata un'analisi qualitativa di alcuni testi tradotti dall'italiano all'inglese e viceversa, ed elaborati da alcuni traduttori automatici presenti sul mercato. Infatti, la Traduzione Automatica è un servizio ormai abbastanza comune e per questo motivo non è più appannaggio di servizi meramente professionali. Il caso più emblematico è probabilmente quello di Systran⁽³⁾, azienda leader del settore del campo dell'Intelligenza Artificiale focalizzata alla traduzione, presente sul mercato da decenni (è stata la prima azienda privata a lanciare un traduttore automatico commerciale, cioè non destinato ad enti governativi). Attualmente, Systran fornisce un *tool* online che permette la Traduzione Automatica

(2) J. SCHÜTZ, B. RIPPLINGER, *Towards Text-Based Machine Translation*, in «Proceedings of the First Conference of the Association for Machine Translation in the Americas», 1994, Columbia, Maryland, USA, pp. 185-194.

(3) <http://www.systran.net/>.

di testi. Questa funzione, basilare nella sua architettura, sfrutta un database ibrido, dunque utilizza la stessa tecnologia dei suoi prodotti *re-tail* acquistabili da professionisti ed aziende. La differenza sta nelle tipologie di applicazione dei due prodotti: nella versione commerciale, il sistema si integra a software di uso comune, come le suite di programmi di scrittura, trattamento testi e fogli di calcolo, sino ai programmi di progettazione di siti web. La versione gratuita e disponibile online mette a disposizione una semplice casella di testo, di per sé intuitiva ma che non lascia molto spazio a personalizzazioni varie (in realtà, essa presenta delle difficoltà per alcune coppie linguistiche per via della presenza di caratteri speciali, tipici di alcune lingue): con una semplice procedura di *copy and paste* è possibile ottenere una traduzione di piccoli testi (la dimensione testuale è un requisito vincolante, si vedrà in seguito) da una *source language* ad una di arrivo. Ovviamente, non tutte le combinazioni linguistiche sono disponibili, tuttavia sono possibili numerose concomitanze: per questo motivo, la qualità traduttiva è un dato non universale, ma che risente della capacità, da parte della macchina, di assimilare le regole grammaticali e lessicali e di trasporle in un'altra lingua con caratteristiche anche molto differenti.

Per questo motivo, il lavoro si concentrerà sull'analisi qualitativa di una sola coppia linguistica (Inglese-Italiano e viceversa), focalizzando l'attenzione soltanto su alcune tipologie testuali ed effettuando un confronto tra "macchine" diverse, che ovviamente produrranno testi di output differenti (al pari di testi prodotti da traduttori diversi). Sarà quindi creato un *framework* di riferimento basandosi sulla resa linguistica dei testi *target* attraverso uno schema di valutazione della precisione traduttiva: il fine ultimo sarà quello dell'individuazione di errori al fine di delineare lo stato dell'arte di alcuni sistemi di Machine Translation.

Al pari delle comuni traduzioni "naturali", anche la valutazione di una traduzione artificiale e meccanica non può prescindere dall'individuazione degli errori più comuni per distinguerne la qualità. Se in aggiunta si pensa che, come detto in precedenza, il livello qualitativo delle traduzioni automatiche non riesce ancora a raggiungere la soglia di performance dell'abilità umana, proprio il fattore della non corrispondenza traduttiva diventa la caratteristica "distintiva" per la determinazione

del valore effettivo di questo tipo di testi, e ciò permetterebbe anche una sorta di classificazione dei migliori software di Traduzione Automatica.

Come ribadito in precedenza, sono numerosi i fattori che concorrono alla creazione di una traduzione di qualità da parte delle macchine. È evidente che la base è rappresentata proprio dal *source text*: questo infatti deve rispondere a chiari requisiti di precisione non solo stilistica e linguistica, ma anche nel layout grafico. Un testo gestito con un normale software di trattamento testi, infatti, stabilisce alcune regole ortografiche ben precise, delle convenzioni da rispettare per non incorrere in imperfezioni ed errori. Parimenti, una *submission* di un testo da tradurre per mezzo di *Machine Translation* richiede il rispetto delle stesse convenzioni, altrimenti vi è il rischio che la macchina non riesca ad individuare tutti gli elementi linguistici presenti nel testo di partenza e l'algoritmo di decodifica non riesca a sostituire il dato con il suo corrispettivo in memoria, attuando una fase di *skip* del termine che verrà trattato erroneamente come, ad esempio, un nome proprio che non prevede traduzione e viene lasciato nella stessa forma.

Un'altra variabile che concorre al livello di qualità traduttiva delle macchine è data dal *domain* del testo stesso. Vi sono infatti dei testi più propensi ad una traduzione maggiormente precisa ed efficace rispetto ad altri. Il *field* di un testo, quindi, è una caratteristica che, nel quadro del livello di performance della *Machine Translation*, assume un suo peso specifico. Il tema-cardine di un testo, infatti, può determinare in un certo senso la leggibilità e la scorrevolezza di un testo di output: qui entrano in gioco alcuni fattori stilistici e sintattici, nonché, come detto, la specificità terminologica di un testo, caratteristiche queste che individuano proprio il settore del testo di partenza. Un testo che tratta di linguaggi specialistici, infatti, molto probabilmente non incorrerà nell'ambiguità semantica di un termine, poiché caratteristica di testi non generici è proprio quella di evitare la polisemia, favorendo la monoreferenzialità⁽⁴⁾. Questo aspetto in un certo senso favorisce la macchina, la quale non è costretta a dover cercare nel proprio dizionario l'accezione di un termine più consona al contesto (e si vedrà proprio come la mancanza di disambiguazione sia una delle caratteristiche più soggette

(4) Cfr. M. GOTTI, *I linguaggi specialistici*, Caratteristiche linguistiche e Criteri pragmatici, La Nuova Italia, Firenze 1991.

ad errore), per cui i cosiddetti tecnicismi risultano un vantaggio relativo nel campo della *Machine Translation*. Prerequisito fondamentale, tuttavia, resta la presenza del termine nel database o nel dizionario del software di traduzione, che altrimenti, come nel caso descritto in precedenza, si ritroverebbe ad affrontare il termine come un termine proprio o non conosciuto, attivando la fase di *skip* che nella maggior parte dei casi si tramuta in una traduzione inesatta, inefficace o incompleta.

Altra caratteristica legata a questo tipo di testi è la non emotività⁽⁵⁾, vale a dire il tono neutro tipico dei testi specialistici. Al di là dell'associazione di non emotività collegata alla macchina (i computer sono dei dispositivi "freddi", che basano la propria intelligenza su associazioni algoritmiche e non deduttive), è bene notare che anche questa caratteristica può risultare vantaggiosa nell'elaborazione di testi da parte dei software di *Machine Translation*. Un testo neutro avrà uno stile diretto, che sviluppa i concetti in maniera molto logica e consequenziale. Ne deriva che la mancanza di linguaggio enfatico porta ad uno stile di logica proposizionale più orientato verso una modalità epistemica della lingua⁽⁶⁾, modalità più razionale che probabilmente meglio si confà alle caratteristiche stilistiche e logiche della Traduzione Automatica.

Per questi motivi, i criteri di selezione dei testi di partenza da sottoporre a Traduzione Automatica e alla conseguente analisi qualitativa dal punto di vista linguistico includeranno dei testi che trattano temi settoriali, e che utilizzano un linguaggio specifico. Tuttavia, i testi di carattere più colloquiale, informale e generico non possono essere esclusi a priori dalla presente analisi. La selezione di vari testi specifici non solo potrà fornire livelli qualitativi di analisi su basi differenti derivanti da tipologie testuali difformi, ma fornirebbe una casistica più completa e capillare della lingua, rappresentando non solo varietà linguistiche diverse, ma anche livelli performativi comparabili relativi alla *Machine Translation*.

La scelta dei testi, dunque, sarà affidata in base a criteri enucleati nel Capitolo III del lavoro.

(5) Ivi, p. 20.

(6) M.A.K. HALLIDAY, *An introduction to Functional Grammar*, Edward Arnold, London 1985.

INTRODUZIONE

IL RUOLO DELLE MACCHINE E PROBLEMATICHE LINGUISTICHE CORRELATE

Man is still the most extraordinary computer of all.

JOHN KENNEDY, 21 maggio 1963.

1. Il rapporto fra uomo e macchina: evoluzione cronologica

La società contemporanea è caratterizzata da una massiccia presenza delle macchine, in ogni forma e in ogni ambito. Macchine capaci di calcolare, sostituire l'uomo in numerose operazioni, velocizzando le attività e garantendo un livello di precisione altrimenti impossibile. Non passa giorno che l'uomo non ricorra ad artifici non umani per compiere numerose operazioni: si pensi alle automobili, veri e propri "motori" delle economie nazionali, a loro volta — in aggiunta alla forza prettamente meccanica — sempre più corredate di accessori tecnologici e ausili che sono definiti "intelligenti". Dunque, le macchine svolgono un ruolo chiave nella collettività di stati e culture, proprio per via del loro positivo impatto nelle società e del loro avanzato livello di tecnologia.

Il rapporto tra l'uomo e le macchine — definite proprio come «dispositivi o apparecchi costruiti [...] in modo che il moto relativo di questi trasmetta o anche amplifichi la forza umana o animale»⁽¹⁾ — trova senz'altro la sua più evoluta espressione negli ultimi tre secoli (con particolare enfasi nel Novecento e con la conseguente evoluzione che comporterà il nuovo millennio), ma ha radici ben più antiche⁽²⁾. Basti pensare

(1) AA.VV., *Vocabolario della lingua Treccani*, Treccani Editore, Roma 2009.

(2) V. MARCHIS, *Storia delle macchine. Tre millenni di cultura tecnologica*, Laterza Editore, Bari-Roma 2005.

all'aratro, strumento fondamentale per lo sviluppo delle economie agricole medievali, considerato, proprio in virtù della sua capacità di svolgere un lavoro "per conto" dell'uomo (sebbene la fase di trasferimento della forza fosse affidata proprio all'uomo o, in alternativa, agli animali), uno strumento di innovazione tecnologica. Il Rinascimento è un fiorire di macchine, soprattutto in campo edilizio: di fondamentale importanza sono le intuizioni di Leonardo da Vinci, che nei suoi *Codici*⁽³⁾ ha lasciato un'importantissima eredità fatta di schemi ingegneristici all'avanguardia, tra cui i numerosi strumenti di trasmissione del movimento (anche nel campo dell'idraulica) o la cosiddetta "macchina per volare", un dispositivo che riusciva a tramutare in realtà il *Mito di Icaro*⁽⁴⁾, nonché precursore e principio primo della ben più moderna aviazione.

È dal Settecento che si cominciano ad intravedere le potenzialità vere e proprie delle macchine. Il secolo per eccellenza dell'Illuminismo, ossia dell'esaltazione massima della Ragione dell'uomo, prospetta un'evoluzione dei mezzi di produzione che porterà alla coniazione di un concetto basilare della storia moderna: la cosiddetta rivoluzione industriale. È la macchina la protagonista assoluta di tale cambiamento: nel suo primo periodo — definito appunto prima rivoluzione industriale — si diffonde il concetto di industria su larga scala, soprattutto nel settore tessile e metallurgico, grazie all'introduzione di macchinari che rendevano il lavoro umano più semplice, agevole e veloce. La protagonista di questa fase è la macchina a vapore (definibile come lo spartiacque tra il concetto di macchina più "classico" e quello moderno), vale a dire un motore in grado di velocizzare il settore produttivo e quello dei trasporti, facilitando di gran lunga il lavoro umano. Con l'avvento dell'elettricità — il simbolo della seconda rivoluzione industriale — il processo di interazione tra l'uomo e la macchina si fa sempre più stretto, e il treno prima, l'automobile poi, diventano gli emblemi di tale rivoluzione, poiché in grado di muoversi attraverso la conduzione dettata dall'uomo stesso.

Ma è il Novecento, come descritto in precedenza, a rappresentare il "Secolo delle macchine". L'importanza di questi strumenti si attesta su scala globale, e viene riconosciuta in tutti i campi della cultura: basti

(3) L. DA VINCI, *Codice del volo [Codex of flight]: dagli uccelli alla macchina per volare*, ed. curata da Edoardo Zanon, Ed. Leonardo3, Milano 2007.

(4) P. OVIDIO NASONE, *Le Metamorfosi*, ed. curata da Giovanni Andrea dell'Anguillara, Giuseppe Orologi, Francesco Turchi, 1610, p. 122.

pensare al Futurismo, movimento culturale, artistico e sociale che si prefigurava come forma estrema di modernismo e che aveva nella macchina il suo emblema principale. Infatti, nel celebre *Manifesto* redatto da Filippo Tommaso Marinetti nel 1909, oltre all'esaltazione di virtù come «Il coraggio, l'audacia, la ribellione» compare anche “una bellezza nuova”, rappresentata dalla “bellezza della velocità”⁽⁵⁾, e proprio la macchina (l'automobile, ma in genere qualsiasi mezzo di locomozione moderno) era in grado di rappresentare in pieno questo spirito.

La terza rivoluzione industriale è quella che inizia in pieno '900 e giunge sino ai giorni nostri. L'automobile, come detto, è parte integrante delle società moderne, ma se si dovesse scegliere un simbolo che possa rappresentare perfettamente l'era contemporanea, la scelta cadrebbe senza ombra di dubbio sui computer.

L'avvento del computer, o calcolatore, è databile intorno alla prima metà dell'Ottocento: nelle sue forme più rudimentali, si trattava di puri e semplici calcolatori, alimentati a vapore. Una *macchina* in piena regola dunque, a schema tuttavia chiuso, ossia capace soltanto di effettuare calcoli basilari. È il secolo successivo, tuttavia, a segnare l'espansione definitiva dei computer: nella prima metà del secolo, questi elaboratori aumentano la potenza di calcolo, ma continuano a limitarsi ad una funzione prettamente matematica. È con Alan Turing che il concetto di computer prende forma nella sua struttura più moderna: considerato il padre della moderna informatica, Turing ha progettato calcolatori sempre più efficienti e ideato principi divenuti poi fondamentali per lo sviluppo dei computer nel campo di ciò che attualmente chiamiamo *Information Technology* (o IT).

Il computer, nell'accezione corrente, è dunque in grado di calcolare: tuttavia, nuovi obiettivi si aprono all'orizzonte. Il XX secolo si è chiuso con l'esplosione definitiva del concetto di personal computer, ossia la possibilità per tutti di avere a disposizione una macchina da calcolo con potenzialità enormi. Nel 2008 la rivista statunitense *New Scientist* pubblicò una stima del numero di PC presenti nel mondo, dichiarando che in quell'anno fu superata la cifra del miliardo di unità, con un margine di crescita esponenziale, tanto da raggiungere il doppio della cifra entro il

(5) F.T. MARINETTI, *Fondazione e Manifesto del Futurismo*, in “Le Figaro”, Parigi, 20 Febbraio 1909.

2014⁽⁶⁾; in realtà alcune statistiche fermano il dato al di sotto della soglia indicata, con un trend in lieve discesa negli ultimi anni⁽⁷⁾. Tuttavia, sarebbe riduttivo limitarsi al solo numero di questi terminali per poter quantificare le curve di crescita nell’ottica di un’espansione della tecnologia di consumo a livello mondiale. Infatti, il tradizionale strumento diffusosi stabilmente per almeno tre decenni in ambienti domestici, lavorativi e di ricerca non può essere considerato l’unico dispositivo in grado di effettuare operazioni computazionali a vario titolo e per diversi usi. A riprova di ciò, si è assistito ad una massiccia proliferazione di strumenti simili al tradizionale PC in termini di possibilità di gestione applicativi, sebbene sia cambiata la modalità di fruizione in favore di strumenti come i moderni smartphone o forme ibride come i tablet, i quali stanno affiancando — e in alcuni casi, completamente sostituendo — il novero delle opportunità digitali a disposizione degli utenti. Numeri a parte è indubbio, tuttavia, che si sia radicato uno scenario allargato a livello globale: con l’avvento di Internet, è possibile comunicare in tempo reale abbattendo il fattore–distanza, considerato fino a qualche tempo fa un deterrente allo sviluppo. Ecco perché un buon *benchmark* non è rappresentato dal numero di terminali, ma di terminali connessi o di utenti che hanno a disposizione una connessione alla Rete. Il sito *Web Statista*, ad esempio, indica che tale numero ha superato, ad aprile 2022, la soglia di 5 miliardi, pari al 63,1% della popolazione⁽⁸⁾. A questi numeri bisogna aggiungere tutti quei dispositivi ‘indiretti’ che effettuano comunque computazioni digitali e/o connesse per vari fini: si pensi a tutti gli strumenti di infotainment installati all’interno di macchine cosiddette analogiche (come ad esempio le autovetture, la cui prestazione di guida dipende da strumenti elettronici per diagnostica o miglioramento della stabilità o velocità) e che con l’integrazione computazionale sono in grado di interfacciarsi in maniera diretta con l’utente. In questo senso, si parla di macchine o dispositivi smart, e

(6) New Scientist staff + Reuters, *Pc Population reaches a billion as e-waste piles up*, *New Scientist*, Ed. Reed Business, Sutton (UK) 23 Giugno 2008. Link diretto all’articolo sull’edizione online: <http://www.newscientist.com/article/dn14186-pc-population-reaches-a-billion-as-ewaste-piles-up-.html>.

(7) *Installed base of personal computers (PCs) worldwide from 2013 to 2019 (in millions)*, disponibile sul sito Statista.com, all’indirizzo <https://www.statista.com/statistics/610271/worldwide-personal-computers-installed-base/>.

(8) *Worldwide digital population July 2022*, <https://www.statista.com/statistics/617136/digital-population-worldwide/>.

questo prefigura una nuova caratteristica da parte della macchina: quella di interagire a seconda di variabili diverse, e con risultati e decisioni a loro volta differenti e non più limitate. Resta indubbio che nella maggior parte dei casi tali interazioni siano veicolate e a favore dell'utente finale, che resta sempre l'uomo. Ma gli scenari contemporanei aprono la strada all'ultimo, forse definitivo passo della conquista da parte delle macchine. Il traguardo è ambizioso: arrivare a creare una macchina capace non più di calcolare, ma anche di *pensare* in maniera autonoma.

Questo obiettivo è compito di varie discipline, soprattutto dell'*Artificial Intelligence* e della Robotica. La domanda di fondo è: "Può una macchina pensare"⁽⁹⁾? In altre parole: è possibile costruire artificialmente una macchina in grado di riprodurre gli stessi passaggi logici della mente umana? Questa riflessione segue una teoria elaborata più di quattro secoli prima da Thomas Hobbes nella sua opera principale, il *Leviatano*. Secondo Hobbes, infatti, la mente umana — e, di conseguenza, la sua capacità di ragionamento — non è altro che il frutto di un calcolo, simile a quello matematico⁽¹⁰⁾. Infatti, al pari delle operazioni aritmetiche, un'addizione di sostantivi porta ad un'affermazione, e a sua volta una somma di affermazioni porta all'enunciazione di un sillogismo, e così via⁽¹¹⁾. La questione della capacità della macchina di replicare in tutte le sfaccettature la logica della mente umana è quindi un tema già affrontato ben prima dell'avvento della moderna elettronica; tuttavia, all'alba del terzo millennio, quest'utopia sviluppata in era illuministica sta diventando pian piano realtà. E allora i passi da gigante compiuti dall'industria della robotica fanno comprendere che le macchine sono in grado di replicare (seppure in maniera ancora non identica all'uomo) i procedimenti logici del cervello umano che, se confrontato con queste macchine con un principio di razionalità nel loro *core*, appare davvero come una macchina perfetta⁽¹²⁾, velocissima nella trasmissione delle informazioni e modello perfetto di imitazione artificiale.

(9) A.M. TURING, *Computing machinery and intelligence*, pp. 121–157, trad. italiana in Lolli G. (a cura di), «Macchine calcolatrici e intelligenza, Intelligenza meccanica», Bollati Boringheri, Torino 1994.

(10) T. HOBBS, *Leviathan*, first published 1651, republished 2008 by Forgotten Books, 2008, p. 19.

(11) Ivi, p. 24.

(12) H. MORAVEC, *When will computer hardware match the human brain?*, in «Journal of Evolution and Technology», vol. 1, published by the Institute for Ethics and Emerging Technologies, Connecticut USA 1998.

La robotica è, nei suoi campi di applicazione, un settore multi-disciplinare che punta allo sviluppo di una macchina in grado di percepire gli stimoli derivanti dall'ambiente esterno, interagendo con esso. Questa branca della scienza trova risvolti pratici nel campo della medicina (basti pensare alle moderne tecniche di chirurgia operate a distanza, con bracci robotici in grado di effettuare l'intervento a seconda dei comandi ricevuti, o ai più "semplici" ausili nelle sale operatorie), dell'industria (le catene di montaggio automatizzate che accelerano il processo di produzione: si pensi al settore automobilistico, nel quale le fasi di verniciatura, montaggio e saldatura sono effettuate da bracci meccanici-automatizzati), ma anche nel campo della strategia militare (i ben noti droni, aerei senza pilota comandati a distanza ed utili per le ricognizioni aeree in territori nemici) e persino nel campo dell'ingegneria spaziale (la sonda *InSight* prima e il veicolo autonomo *Perseverance*, mandati in missione esplorativa su Marte rappresentano le applicazioni tecnologicamente più avanzate al momento⁽¹³⁾, rappresentando alcune tra le missioni esplorative e preparatorie per il "grande salto" verso l'esplorazione umana del pianeta rosso, prevista sempre più verosimilmente entro i prossimi due decenni).

Dunque: i robot saranno capaci di pensare? Le macchine arriveranno a sostituire in tutto e per tutto l'uomo? In realtà queste tematiche sembrano appartenere più al campo della fantascienza piuttosto che a quello della ben più razionale scienza. Eppure la storia ci insegna che vi sono stati periodi nei quali la spinta tecnologica aveva indotto a pensare che si fosse molto vicini a questo obiettivo. Gli anni '50 del Novecento, ad esempio, sono contraddistinti dall'antologia di racconti di Isaac Asimov dal nome *I, Robot*⁽¹⁴⁾, che ha segnato un'epoca e che ancora continua ad avere proseliti (si pensi ad esempio alla pellicola prodotta nel 2004 e diretta da Alex Proyas che porta lo stesso nome) a distanza di alcuni decenni. Robot capaci di camminare, di correre, di compiere in tutto e per tutto le azioni degli esseri umani. Capaci anche di parlare e di pensare.

Traendo spunto da queste ultime considerazioni, uno degli obiettivi forse più ambiziosi nel campo dell'innovazione tecnologica applicata all'Intelligenza Artificiale è legato proprio ai fenomeni linguistici⁽¹⁵⁾.

(13) <https://mars.nasa.gov/insight/> e <https://mars.nasa.gov/mars2020/>.

(14) I. ASIMOV, *I Robot*, Doubleday Book Club Ed. 1950.

(15) Cfr. P. KUGEL, *Toward a theory of intelligence*, in «Theoretical Computer Science», vol. 317, issues 1-3, Elsevier B.V. 2003.

E tra le varie tipologie associabili al campo dell'NLP (*Natural Language Processing*), ossia quella branca della *Computer Science* e della linguistica che mira all'interazione quanto più accurata possibile tra computer e linguaggio umano (non si confonda il concetto con il linguaggio di programmazione tipico dei computer e indispensabile per l'interazione stessa), menzione speciale merita la riproduzione inter-linguistica di testi. In altre parole, l'obiettivo è quello di permettere alle macchine di tradurre⁽¹⁶⁾.

11. Una macchina in grado di tradurre: la *Machine Translation*

La *Machine Translation* è oggetto di studi da alcuni decenni, e ha conosciuto fasi alterne di sviluppo sin dai suoi albori, databili intorno al terzo decennio del Novecento. La fase apicale degli studi sulla traduzione effettuata dalle macchine è senza dubbio legata a fattori storici e sociali, allorché contingenze derivate dalla situazione politica creatasi all'alba del secondo dopoguerra portarono le due superpotenze mondiali a scatenare una vera e propria guerra a colpi di innovazioni tecnologiche, che ben si identificano con il periodo definito come “la corsa alla Luna”⁽¹⁷⁾. La necessità di reperire informazioni per sottrarre alla controparte avversaria conoscenze nuove in ambito scientifico portò alla creazione di alcuni algoritmi che permisero ai calcolatori dell'epoca — già di loro protagonisti della ricerca in altri ambiti scientifici, primo tra tutti l'ingegneria aerospaziale — di riportare delle informazioni da una *source language* in una *target language* nel minor tempo possibile, sostituendo in maniera meccanica de facto l'intervento umano, poiché la tipologia di testi da tradurre in quel determinato campo portava ad una certa ripetitività che alla lunga sarebbe risultata congeniale alla meccanica di calcolo algoritmica dei *mainframe*. Tale aspirazione derivava comunque da studi iniziati qualche anno prima, ma rimasti allo stadio embrionale proprio per la mancanza di un calcolatore di tipo elettronico che potesse effettuare determinati tipi di operazioni. Tuttavia, il

(16) A tal proposito, Diana Santos considera la *Machine Translation* “the ultimate application of NLP”. Cfr. D. SANTOS, *Toward Language-dependent applications*, in «Machine Translation», vol. 14, issue 2, Springer 1999, p. 104.

(17) P. MAGIONAMI, *Gli anni della luna. 1950–1972: l'epoca d'oro della corsa allo spazio*, Springer, Milano 2009, p. 34.

dopoguerra per gli Stati Uniti e per l'Unione Sovietica rappresentò il tempo della definitiva espansione: l'uomo di lì a poco avrebbe varcato i confini terrestri, per cui qualsiasi traguardo non era considerato insuperabile. E, restando sulla Terra, l'obiettivo ambizioso era proprio quello di far tradurre le macchine con la stessa qualità della performance umana, con il vantaggio di effettuarla in un lasso di tempo nettamente inferiore.

I primi studi furono incoraggianti, poiché gli interventi di pura sostituzione lessicale portarono a un livello qualitativo senz'altro superiore rispetto alle aspettative iniziali: i margini di crescita lasciavano presagire una risoluzione del problema entro breve tempo. Tuttavia, si scoprì che il sogno del “test di Turing” era ben lungi dall'essere completato ed affinato. Gli studi figli della Guerra Fredda, di fatto, si scontrarono a metà degli anni Sessanta con un report⁽¹⁸⁾ di una Commissione americana che segnò il tramonto della prima fase della *Machine Translation*, allorché si determinò lo stallo dello *state of art* nel campo della Traduzione Automatica.

Tuttavia, un obiettivo così ambizioso non poteva conoscere una definitiva interruzione. Gli scenari aperti dalla diffusione a livello globale dei moderni computer portarono ad una sostanziale modifica degli scenari geografici. Se dunque la prima fase degli studi nel campo della Traduzione Automatica vede protagoniste le due superpotenze mondiali, ossia gli Stati Uniti e l'Unione Sovietica, la fase successiva — il cui inizio è databile intorno ai primi anni degli anni '70 — vede protagoniste nazioni nuove, come il Giappone, il Canada e la neonata Comunità Economica Europea (CEE)⁽¹⁹⁾. La globalizzazione, dunque, ha aperto e apre tuttora nuovi scenari nel campo della *Machine Translation*, poiché le capacità sempre maggiori di calcolo da parte dei computer, nonché nuove tecniche e nuovi approcci relativi alla problematica della resa interlinguistica sono in fase di sviluppo, e prefigurano ulteriori progressi e miglie in questo campo scientifico. Sembra doveroso tuttavia precisare che, ad oggi, il traduttore umano potrà essere

(18) *Language and machines: computers in translation and linguistics. A report by the Automatic Language Processing Advisory Committee, Division of Behavioral Sciences, National Academy of Sciences, National Research Council, cit.*

(19) European Commission, *Statistical Portrait of the European Union 2007 – 50 Years of the Treaty of Rome establishing the European Economic Community*, Eurostat 2007, p. 8.