

# Computabilità e traduzione interlinguistica: considerazioni sull'impiego dell'Intelligenza Artificiale in ambito traduttivo

**Marika Angelucci**

Università "G. D'Annunzio" – Chieti-Pescara

*Il presente articolo si inserisce nell'acceso dibattito attorno al settore multidisciplinare dell'Intelligenza Artificiale assumendo una postura specifica: il punto di vista traduttologico. La Traduzione Automatica, infatti, ha costituito la prima applicazione del computer in ambito umanistico, sia in virtù del generico interesse per lo studio del linguaggio umano, data la sua forte connotazione identitaria, sia per via della sua alta formalizzazione. Attraverso una panoramica delle architetture che si sono susseguite, s'intende focalizzare l'attenzione sul proficuo dialogo tra pratica e teoria, a cui si devono le riflessioni moderne sui meccanismi della traduzione, processo decisionale e di negoziazione. Tra queste: i limiti del passaggio interlinguistico, l'entropia dell'informazione, l'anisomorfismo delle lingue naturali, la mutua influenza di linguaggio e cognizione, la rilevanza delle informazioni extra-testuali e degli orizzonti cognitivi dei parlanti, il concetto di fedeltà testuale. L'approfondimento di tali aspetti apre a delle riflessioni di respiro più ampio, come la questione spinosa del rapporto tra intelligenza naturale e intelligenza artificiale, l'essenza della comprensione e il possibile impedimento euristico costituito dall'imitazione.*

## Introduzione

**M**anifestazione iper-moderna dell'aspirazione umana all'autocoscienza<sup>1</sup>, il settore di ricerca interdisciplinare dell'Intelligenza Artificiale (IA) indaga le teorie e i metodi che consentono di comprendere, e quindi di costruire, dei modelli di razionalità, *imitativi* o meno del pensiero e dell'azione umani<sup>2</sup>.

Nell'arco dell'ultimo biennio, soprattutto in seguito al rilascio della chatbot ChatGPT da parte dell'azienda californiana OpenAI, i dispositivi di IA hanno acce-

<sup>1</sup> Arnold Gehlen, *L'uomo nell'era della tecnica*, Roma: Armando, 2003 p. 40, in: Vallori Rasini, *L'essere umano. Percorsi di antropologia filosofica contemporanea*, Roma: Carocci, 2008, p. 117: «angustiato dall'enigma della sua esistenza e della sua stessa essenza, l'uomo non ha altra risorsa che cercare di interpretarsi passando attraverso un non io, attraverso qualcosa di diverso dall'umano».

<sup>2</sup> Peter Norvig – Stuart J Russel, *Artificial Intelligence: A Modern Approach*, Londra: Prentice Hall, 1994.

so il dibattito pubblico. Anche tra i non specialisti emergono curiosità riguardo il loro funzionamento, scetticismo o entusiasmo in merito alla possibilità di farne uso in numerosi ambiti della vita quotidiana e lavorativa e dubbi sulle politiche nazionali e internazionali necessarie a regolarne un utilizzo ragionevole e sicuro. Il presente articolo si inserisce nel dibattito attorno all'Intelligenza Artificiale, con particolare attenzione alla possibilità di sfruttare dei dispositivi per portare a termine il difficile compito della traduzione interlinguistica<sup>3</sup>.

Dalla nascita programmata, in occasione del celebre convegno del Dartmouth College (New Hampshire) del 1956, il settore dell'IA unifica gli sforzi delle numerose discipline – la filosofia, la matematica, la psicologia, la cibernetica, l'informatica e la linguistica – che, adottando criteri e metodi specifici, hanno perseguito individualmente un obiettivo condiviso: la modellizzazione<sup>4</sup> dell'intelligenza umana e, più in generale, della razionalità.

Tra le peculiarità dell'intelletto umano che ci si è proposti di ricreare attraverso degli strumenti automatici, il linguaggio, percepito come elemento fortemente identitario dell'uomo, ha da subito rivestito un ruolo chiave<sup>5</sup>. Proprio in virtù della sua predisposizione alla formalizzazione, non è strano che «il primo progetto di applicazione umanistica del neonato computer»<sup>6</sup> sia stata proprio la Traduzione Automatica<sup>7</sup>: il ricorso a processi automatici al fine di trasportare i significati da un testo di partenza  $S$ , in cui sono codificati nella lingua  $c_s$ , a un testo d'arrivo  $T$ , in cui sono ricodificati nella lingua  $c_t$ .

Se il progetto di utilizzare il computer per la traduzione interlinguistica è stato il primo fra quelli concepiti al di fuori dell'ambito matematico, tuttavia, ciò è avvenuto in assenza dei presupposti teorici che, secondo la prassi tradizionale, avrebbero dovuto orientare la pratica: «l'analisi del concetto e dei metodi della traduzione sarà sollecitata proprio dagli sforzi per automatizzarla»<sup>8</sup>. Alla base di questo impegno, oltre agli interessi puramente intellettuali, vanno rintracciate delle importanti forze storiche: dei presupposti strategici, legati alle dinamiche di spionaggio del secondo dopoguerra, e dei presupposti economici. La possibilità di affidare la traduzione a una macchina, infatti, si riflette nell'economicità dell'attività traduttiva, tanto in

<sup>3</sup> Si è scelto di non approfondire in questa sede le numerose e complesse questioni etiche connesse al settore dell'IA, che meriterebbero un'ampia trattazione a parte.

<sup>4</sup> Tito Orlandi, *Informatica testuale. Teoria e prassi*, Roma: Editori Laterza, 2010, p. 155: «Il modello è un'entità,  $M$ , concreta o astratta, che abbia la proprietà di poter essere utilizzata per simulare e, dunque, entro certi limiti, spiegare, comprendere, il comportamento di un'entità,  $P$ , il prototipo, anch'essa concreta o astratta».

<sup>5</sup> Per approfondimenti: Giosuè Baggio, *Neurolinguistics*, Cambridge: The MIT Press, 2022; Robert C. Berwick – Noam Chomsky, *Perché solo noi. Linguaggio ed evoluzione*, Torino: Bollati Boringhieri, 2016; Claudia Casadio, *Logica Informazione Automazione*, Corfino: primeVie edizioni, 2012.

<sup>6</sup> T. Orlandi, *Informatica testuale. Teoria e prassi*, cit., p. 142.

<sup>7</sup> Nell'articolo il termine occorre in due forme diverse: in maiuscolo quando lo si usa per indicare il settore, in minuscolo quando invece indica il prodotto della codifica.

<sup>8</sup> T. Orlandi, *Informatica testuale. Teoria e prassi*, cit., p. 142.

termini di costi della prestazione, quanto di tempo di esecuzione.

Per quanto concerne il motore intellettuale della ricerca traduttologica, che in questa sede riveste il ruolo di maggiore interesse, è importante notare che, come emergerà dall'exkursus storico che segue, solo grazie alla costruzione di alcune architetture e all'osservazione dei difetti di volta in volta riscontrati nell'elaborazione dei dati è stato possibile far progredire la scienza dietro la tecnica<sup>9</sup>. Attraverso la sperimentazione, per tentativi ed errori, è stato possibile perfezionare la conoscenza dei meccanismi che ci si proponeva al contempo di modellizzare: ogni architettura proposta metteva in risalto nuove problematiche prima ignorate, permettendo così di approfondire la conoscenza e, conseguentemente, di mettere a punto nuovi modelli che conducessero a un risultato sempre più adeguato e soddisfacente.

### L'insufficienza della traduzione diretta

La storia della Traduzione Automatica ha inizio nel 1949, quando Warren Weaver e Claude Shannon, matematici statunitensi, applicano il modello crittografico alla traduzione, da loro considerata un semplice processo di decifrazione del codice linguistico: il messaggio è codificato nella lingua di partenza e, per essere compreso dal ricevente, dev'essere individuato e tradotto, ossia rispettivamente decodificato e ricodificato nella lingua d'arrivo.

Sulla base di questa concezione, la prima architettura realizzata è stata quella "di traduzione diretta": il sistema si occupava di trovare l'equivalenza formale tra i singoli elementi della lingua di partenza e della lingua d'arrivo, sfruttando un dizionario bilingue. Certamente la traduzione *mot-à-mot* realizzata da una simile architettura non poteva essere soddisfacente: si trattava di un approccio semplicistico, fondato su una congruenza formale di fatto insussistente tra le lingue naturali<sup>10</sup>. Proprio grazie ai limiti osservati è stato possibile ripensare la pratica traduttiva e riflettere approfonditamente su questioni spinose come l'anisomorfismo linguistico, l'onnipotenza semantica e la non-linearità del codice linguistico.

Per questi motivi, l'equivalenza tra il testo di partenza e il testo d'arrivo è ottenibile a livello contenutistico, per via della capacità, condivisa dai codificatori, di razionalizzare, e quindi codificare, i mondi (soggettivamente) percepiti, non a livello formale. In altri termini, visto il principio di onnipotenza semantica, le lingue «differiscono essenzialmente per ciò che *devono* esprimere, non per ciò che *possono* esprimere»<sup>11</sup>. Ogni lingua organizza in un modo che le è proprio un certo spazio di significato, reagendo all'equilibrio con la visione del mondo del parlante, e dun-

<sup>9</sup> Thierry Poibeau, *Machine Translation*, Cambridge: The MIT Press, 2017.

<sup>10</sup> Per approfondimenti: Noam Chomsky – Andrea Moro, *I segreti delle parole*, Milano: La nave di Teseo, 2022; Andrea Moro, *Le lingue impossibili*, Milano: Raffaello Cortina, 2017.

<sup>11</sup> Roman Jakobson, *Aspetti linguistici della traduzione*, in: *Teorie contemporanee della traduzione*, a cura di S. Nergaard, Torino: Bompiani, 2014 [1995].

que evidenziando certi aspetti della realtà piuttosto che altri, cogliendone alcuni elementi piuttosto che altri: le lingue naturali sono non-congruenti, non-sovrapponibili e non-lineari.

Di conseguenza, benché il tasso di fungibilità delle diverse strategie traduttive sia variabile anche in ragione della natura del testo che s'intende ricodificare, generalmente il ricorso alle "operazioni di equivalenza diretta"<sup>12</sup> – ossia trascrizione, prestito, calco, traduzione letterale: strategie traduttive che restano confinate alla superficie testuale – risulta inferiore rispetto alle "operazioni di equivalenza indiretta", operazioni oblique che prevedono una manipolazione informativa ben più marcata, ponendosi «all'insegna della negoziazione»<sup>13</sup>.

Per via dell'irriducibilità dei sistemi linguistici, la letteralità è un fenomeno raro: l'equivalenza globale tra i testi è garantita da operazioni asimmetriche – come la trasposizione, la modulazione, l'adattamento –, risultato dell'intervento plastico del traduttore. La traduzione è un "processo decisionale"<sup>14</sup> e, sulla base della gerarchizzazione degli elementi informativi, il testo d'arrivo viene modellato attorno a un nucleo di presunta fedeltà<sup>15</sup>, con lo scopo di minimizzarne l'entropia, che pure resta ineliminabile:

«Translations are inevitably partial: meaning in a text is always overdetermined, and the information in a source text is therefore always more extensive than translation can convey. Conversely, the receptor language and culture entail obligatory features that shape the possible interpretations of the translation [...]»<sup>16</sup>.

In altri termini, T dirà *quasi* la stessa cosa di S, sia in virtù di quanto il soggetto codificatore di S concepisce ed esprime attraverso la lingua  $c_s$ , sia in virtù di quanto il soggetto codificatore di T recepisce sulla base della sua idea di mondo e ricrea attraverso la lingua  $c_t$ : i rispettivi orizzonti cognitivi costituiscono i sistemi di riferimento all'interno dei quali il testo prende forma<sup>17</sup>.

<sup>12</sup> Josaine Podeur, *Jeux de traduction / Giochi di traduzione*, Napoli: Liguori editore, 2008, p. 27-52.

<sup>13</sup> Umberto Eco, *Dire quasi la stessa cosa*, Torino: Giunti Editore, 2003, p. 16-17.

<sup>14</sup> Jiří Levý, *La traduzione come processo decisionale*, in: *Teorie contemporanee della traduzione*, a cura di S. Nergaard, Torino: Bompiani, 2014 [1995].

<sup>15</sup> Umberto Eco, *Riflessioni teorico-pratiche sulla traduzione*, in: *Teorie contemporanee della traduzione*, a cura di S. Nergaard, Torino: Bompiani, 2014 [1995]; U. Eco, *Dire quasi la stessa cosa*, cit.

<sup>16</sup> *Translation and Power*, ed. by M. Tymoczko, E. Gentzler, Amherst and Boston: University of Massachusetts Press, 2002, p. XVIII; e Susan Bassnett, *Translation*, New York: Routledge, 2014, p. 169.

<sup>17</sup> Per approfondimenti: Antoine Berman, *La traduzione e la lettera o l'albergo nella lontananza*, Macerata: Quodlibet, 2003; Raffaella Bertazzoli, *La traduzione: teorie e metodi*, Roma: Carocci, 2019 [2006]; *Teorie contemporanee della traduzione*, a cura di S. Nergaard, Torino: Bompiani, 2014 [1995]; *Terminologia della traduzione*, a cura di M. Ulrych, Milano: Hoepli, 2002.

## Le architetture di traduzione successive

Constatata l'insufficienza dell'architettura di traduzione diretta, il passo successivo verso una traduzione automatica migliore ha riguardato l'analisi preliminare del testo di partenza, filtrato e organizzato mediante l'applicazione di regole linguistiche. Nell'architettura *rule-based* emergono l'influenza delle teorie generativo-trasformazionali a essa contemporanee e la relativa presa di coscienza della struttura gerarchica del linguaggio<sup>18</sup>. Nonostante l'evidente avanzamento rispetto al passato, anche nel nuovo modello persistevano dei limiti traduttivi importanti, derivanti soprattutto dalla ristrettezza del ventaglio di informazioni di volta in volta considerate nella computazione. Il sistema, infatti, non solo continuava a lavorare su singoli elementi, ma soprattutto non aveva accesso né ai dati di co-testo, né tantomeno ai dati di contesto, irrinunciabili per la buona riuscita della traduzione, visto che nella maggior parte dei casi essi costituiscono le uniche fonti di disambiguazione strutturale e semantica.

La nascita di Internet ha rappresentato un importante punto di snodo anche nel settore traduttologico perché ha fornito ai ricercatori un enorme database non strutturato di cui servirsi. Dal confronto con la rete deriva l'architettura *corpus-based*, fondata sull'impiego dei corpora paralleli e sulla tecnica dell'*allineamento*.

Disponendo degli allineamenti di testo e traduzione, tipicamente per frase, la macchina può processare le informazioni in due modi principali: può memorizzare le equivalenze in vista delle traduzioni future e disporne in maniera cristallizzata – è il caso della *example-based translation* –, oppure può processare le corrispondenze di cui dispone per creare dei modelli statistici che orientino in modo meno prescrittivo le traduzioni future – in questo caso si parla di *Statistical Machine Translation*, SMT.

Il secondo caso, ossia l'applicazione dei metodi statistici alla traduzione, ha segnato il vero punto di svolta della Traduzione Automatica perché ha prodotto la cosiddetta "equazione fondamentale", madre di tutti i modelli traduttivi successivi:  $T = \operatorname{argmax}(T) [\operatorname{Pr}(S|T) * \operatorname{Pr}(T)]$ <sup>19</sup>.

Questa formula matematica condensa in pochi segni tutto lo stato dell'arte: come simboleggiato dalla funzione *argmax* – che esprime la ricerca dell'elemento del Dominio della funzione la cui immagine è massima – tradurre, in termini di computazione, significa trovare, dato un testo di partenza *S*, il migliore – cioè più probabile – testo d'arrivo *T*. Nel dettaglio, la probabilità del testo di partenza dato il testo di arrivo,  $\operatorname{Pr}(S|T)$ , rappresenta il "modello traduttivo" (*translation model*), un sistema che conosce le possibili traduzioni di unità della lingua *S* in lingua *T* e che, sfruttando dei corpora paralleli, misura la probabilità che la sequenza di parole *S* corrisponda alla sequenza di parole *T*; la probabilità di *T*,  $\operatorname{Pr}(T)$ , rappresenta il "mo-

<sup>18</sup> Per approfondimenti: Noam Chomsky, *Il linguaggio e la mente*, Torino: Bollati Boringhieri, 2010 [2006]; Steven Pinker, *Words and rules*, New York: Basic Books, 2015 [1999].

<sup>19</sup> Thierry Poibeau, *Machine Translation*, Cambridge: The MIT Press, 2017, p. 126-131.

dello linguistico” (*language model*), un sistema che sfrutta dei corpora monolingui e misura la probabilità che T sia una sequenza di parole valida in base alle regole proprie della lingua d’arrivo. L’equazione fondamentale ha, quindi, complessivamente il significato seguente: la traduzione migliore (T’) è quella per cui il “modello traduttivo” e il “modello linguistico” risultano ottimizzati.

Come si nota dall’unità tipica di allineamento, a differenza delle architetture precedenti la SMT consente di lavorare alla traduzione gestendo le informazioni relative alla frase intera e non più solo a singoli elementi. Tuttavia, anche in questa tipologia di architettura persistono delle debolezze legate tanto alla natura dei corpora che, per quanto estesi, sono strutturalmente lontani dall’esaustività delle combinazioni linguistiche, quanto alla perenne inaccessibilità al contesto e al co-testo, se non di immediata adiacenza al segmento testuale considerato.

Conseguentemente, il passo successivo è stato mosso proprio verso una migliore integrazione nel modello delle informazioni co-testuali e contestuali necessarie alla decodifica del significato del testo di partenza e, quindi, alla formulazione di un output che rispettasse i criteri di testualità, primo fra tutti la coerenza.

Questa, per definizione, non è rintracciabile nella forma materiale del testo, non è connessa al testo nella sua dimensione concreta, ma è il risultato di un processo, dell’interazione tra testo e soggetto: il soggetto interpretante, ponendosi in maniera attiva di fronte al testo, ne inferisce il significato facendolo reagire con le informazioni extratestuali in suo possesso, stagliandolo sul suo *background knowledge*<sup>20</sup>.

Affinché la coerenza venga rispettata, si rende necessario integrare nel modello le informazioni extratestuali utili affinché il testo possa essere codificato e ricodificato correttamente. Viste la plasticità e la dinamicità delle idee e del sapere, sarebbe stato poco funzionale cercare la soluzione nella creazione di un database: per quanto esteso, la sua staticità lo avrebbe reso di rapida obsolescenza; pertanto, si è pensato di focalizzare l’attenzione non sulla quantità dei contenuti accessibili, ma sul sostrato e sui meccanismi grazie ai quali tali contenuti vengono generati nella mente umana.

## Le reti neurali

Sono questi i presupposti della *Neural Machine Translation* (NMT), basata sull’architettura della “rete neurale” e sul meccanismo fondamentale dell’*apprendimento*. Le reti neurali sono dei sistemi computazionali complessi d’impianto connessionista, concepiti sin dagli albori del settore dell’Intelligenza Artificiale, ma riscoperti e utilizzati sistematicamente solo a partire dal 2010, momento in cui il resto dell’apparato tecnologico ha potuto soddisfare la richiesta di una simile potenza di calcolo.

<sup>20</sup> Maria E. Conte, *Condizioni di coerenza. Ricerche di linguistica testuale*, Alessandria: Edizioni dell’Orso, 1999.

lo. Il modello della rete neurale si ispira alla struttura fisiologica comunemente ritenuta essere il sostrato dell'intelligenza naturale: il *connettoma*, l'insieme delle connessioni tra i neuroni del sistema nervoso, la «mappa di navigazione del nostro cervello»<sup>21</sup>. Composta da nodi e vettori, la rete neurale modella i pirofori e gli assoni biologici: i nodi rappresentano dei concetti e i vettori i mezzi di trasmissione dell'impulso. A ogni vettore è associato un peso (*weight*) che esprime quantitativamente l'efficacia sinaptica della connessione tra i nodi interessati: più un vettore è "pesante", più quella connessione è forte e, quindi, probabile e stabile.

Il percorso neurale più idoneo all'elaborazione dell'informazione viene stabilizzato grazie al processo di *apprendimento*: poiché è improbabile che le stime iniziali della macchina siano soddisfacenti, è necessario che questa venga allenata (nella fase di *training*) a processare le informazioni secondo le aspettative dell'utilizzatore; per ottenere dei risultati conformi, alla macchina vengono sottoposti numerosi esempi da elaborare, su cui l'operatore interviene con delle correzioni. Sulla base di tale *editing*, la macchina ripercorre a ritroso il percorso individuato dai nodi, rivede il peso dei vettori e stabilisce il cablaggio che conduce all'output migliore, ossia all'ottimizzazione dei valori numerici.

Come nel caso dell'architettura *corpus-based*, la *Neural Machine Translation* fa ricorso ai corpora paralleli e all'analisi statistica; tuttavia, diversamente da quanto avveniva con le architetture precedenti, la macchina, a partire da queste stesse risorse, non si limita più a eseguire un'associazione lineare, ma riesce a sfruttare in maniera dinamica le informazioni di cui dispone, creandone una mappa. In effetti, siccome la rappresentazione del testo è affidata a interi gruppi di nodi-concetto, è distribuita su interi strati (*layers*) di unità informative, diventa possibile creare uno schema analitico dei significati e, conseguentemente, veicolare il testo dalla lingua di partenza alla lingua d'arrivo senza quei vincoli formali che tanto costringevano le architetture precedenti.

Il panorama attuale dei dispositivi di IA pensati per il trattamento del linguaggio verbale è dominato dai Large Language Model (LLM), architetture neurali estremamente ampie di tipo generativo: dato un *prompt*, la rete genera automaticamente, in maniera non supervisionata, delle sequenze di parole, concatenate sulla base dei dati statistici acquisiti nella fase di *training*. Se implementate in software per la traduzione, queste architetture ne rappresenterebbero uno strumento di predizione statistica senza pari.

Un esempio tra tutti, GPT-3, la terza versione del *Generative Pretrained Transformer* dell'azienda californiana OpenAI: una rete neurale formata da novantotto *hidden layers*, costruiti su centosettantacinque miliardi di parametri. Si tratta di un modello di linguaggio autoregressivo – la variabile in uscita dipende linear-

<sup>21</sup> Sebastian Seung, *Connettoma, La nuova geografia della mente*, Torino: Codice Edizioni, 2016, p. 8.

mente dai valori delle uscite precedenti – che utilizza il meccanismo dell’apprendimento per generare automaticamente delle sequenze di parole a partire da un input sorgente<sup>22</sup>.

Della stessa azienda si ricordano anche ChatGPT – la chatbot che ha suscitato l’interesse generale, anche dei non specialisti, accendendo di recente il dibattito pubblico sui temi dell’Intelligenza Artificiale e della privacy dei dati<sup>23</sup> – e GPT-4, ampio modello multimodale (da immagine o testo a testo) rilasciato a marzo 2023, che costituisce l’evoluzione *scaled* di GPT-3.

## Potenzialità e limiti della traduzione automatica

Il ricorso alla Traduzione Automatica offre numerosi vantaggi, tra cui:

- la convenienza economica, tanto in termini di costi della prestazione, quanto in termini di tempo di realizzazione dell’elaborato; da cui consegue...
- la democraticità dell’accesso ai contenuti codificati nel testo: tradurre rapidamente e a basso costo significa garantire a un pubblico ampio l’accesso a contenuti altrimenti non facilmente fruibili;
- la numerosità delle combinazioni linguistiche a disposizione: la quantità delle combinazioni linguistiche, attive e passive, coperte da un singolo software, e quindi la quantità di testi che un singolo software è in grado di gestire ed elaborare, supera di gran lunga le potenziali capacità di un traduttore umano.

Se i vantaggi appaiono chiari e ben determinati, il profilo dei limiti della Traduzione Automatica si delinea meno nettamente, perché la loro valutazione implica delle riflessioni e delle argomentazioni plurivoche.

Al di là delle problematiche proprie della traduzione *tout court*, legate alle dinamiche della negoziazione dei significati e della dispersione entropica dell’informazione, la macchina deve misurarsi con ulteriori difficoltà, che derivano dal confronto tra la sua struttura costitutiva, che ne determina il funzionamento, e il linguaggio umano, tra cui:

<sup>22</sup> Per approfondimenti: Fabio Ciotti, *I Large Language Model e i “problemi difficili”*, «Informatica Umanistica e Cultura Digitale: il blog dell’ AIUCD», 22 agosto 2023; Luciano Floridi – Massimo Chiriatti, *GPT-3: Its Nature, Scope, Limits, and Consequences*, «Minds and Machines», 30 (2020), p. 681-694; Katherine Elkins – Jon Chun, *Can GPT-3 Pass a Writer’s Turing Test?*, «Journal of Cultural Analytics», 5 (2020), n. 2, <<https://doi.org/10.22148/001c.17212>>; Will D. Heaven, *OpenAI’s new language generator GPT-3 is shockingly good—and completely mindless*, «MIT Technology Review», 20 luglio 2020; Gary Marcus, *The New Science of AI Intelligence*, The Road to AI We Can Trust, 14 maggio 2022; Justin Weinberg, *Philosophers On GPT-3 (updated with replies by GPT-3)*, «Daily Nous», 30 luglio 2020.

<sup>23</sup> Per approfondimenti: Greg Brockman, *The Inside Story of ChatGPT’s Astonishing Potential*, TED Talk, 20 aprile 2023; Gabriel A. Silva, *The ChatGPT Debate: Are We Intelligent Enough To Understand ‘Intelligence’?*, «Forbes», 14 marzo 2023; Noam Chomsky, *The False Promise of ChatGPT*, «The New York Times», 8 marzo 2023; Riccardo Staglianò, *ChatGPT, meraviglie e pericoli dell’intelligenza artificiale*, «Repubblica», 19 aprile 2023; Alessia Conzonato, *ChatGPT è diventata “più stupida”? Le teorie degli utenti sulla regressione*, «Corriere della Sera», 30 luglio 2023.

- l'imprevedibilità data dalla creatività linguistica, una caratteristica peculiare e pervasiva del linguaggio umano diametralmente opposta ai requisiti di esplicitazione e di alta formalizzazione necessari al corretto funzionamento della macchina;
- la mancanza di senso comune, che mina la possibilità stessa di rispettare la coerenza testuale.

In generale, poiché tradurre significa ricercare l'identità di senso e di registro tra un testo di partenza e un testo di arrivo, con il passaggio tanto dei semi inerenti, quanto dei semi afferenti, nonché, nei limiti del possibile, della parte di senso veicolata dalla forma, la creatività linguistica, nel caso della traduzione, diventa uno strumento irrinunciabile per la gestione del livello di entropia. La necessità di svincolarsi dalla forma del testo di partenza per veicolare l'insieme dei significati in un testo d'arrivo che si inserisca correttamente nella prospettiva cognitiva e culturale del ricevente fa della creatività linguistica un requisito indispensabile per la buona riuscita della traduzione.

Concentrando l'attenzione sui fenomeni individuali, un caso particolarmente significativo è rappresentato dalle espressioni metaforiche, figure del linguaggio estremamente complesse e altrettanto frequenti, la cui peculiarità è l'inconsistenza semantica: creano un ponte tra due ambiti semantici separati, anche molto distanti, che condividono un solo seme di significato. Ancor prima che retorico, il ruolo rivestito dalle espressioni metaforiche è cognitivo: sono strategie distintive del pensiero umano che svolgono una funzione euristica. Elena Semino, linguista angloitaliana, spiega, infatti, che la figura retorica concreta non è altro che la materializzazione linguistica di un meccanismo mentale, detto "metafora concettuale", che permette di spiegare qualcosa di nuovo e astratto attraverso il concreto, ampiamente noto<sup>24</sup>. Se la metafora prevede un salto tra campi semantici, tanto più lungo quanto più l'espressione è ardita, evocativa ed efficace, allora, vista la sua scarsa predicibilità, rappresenta una vera sfida per la formalizzazione. Al contempo, però, il problema della codifica dello scarsamente prevedibile si pone con urgenza: essendo uno strumento cognitivo irrinunciabile, "di cui viviamo"<sup>25</sup>, la metafora si manifesta linguisticamente molto di frequente e in ogni tipologia testuale.

Accanto al limite della creatività, un altro ostacolo materiale alla traduzione automatica ottimale è la mancanza di accesso al senso comune. Il *commonsense reasoning* è un altro esempio di meccanismo primario della mente umana, che impara a memorizzare gli andamenti, le relazioni di causa-effetto, i fenomeni tipici della realtà che abita e a tenere in considerazione i dati così inferiti nelle proprie valutazioni. Senza la conoscenza di come funziona il mondo e delle situazioni che vi hanno tipicamente luogo per un corpo che lo abita, spesso diventa difficile rispettare il criterio fondamentale di testualità, ovvero la coerenza.

<sup>24</sup> Elena Semino, *Metaphor in Discourse*, Cambridge: Cambridge University Press, 2008.

<sup>25</sup> George Lakoff – Mark Johnson, *Metaphors We Live By*, Chicago: University of Chicago Press, 2003.

Proprio con la volontà di colmare questa lacuna, Hector Levesque, docente di Computer Science presso l'Università di Toronto, ha proposto di valutare l'intelligenza della macchina attraverso un quesito a scelta binaria la cui soluzione richiede l'accesso a informazioni extratestuali che l'essere umano considera del tutto ovvie, ma che tali non sono per un calcolatore, per cui tutto dev'essere esplicito e correttamente formalizzato<sup>26</sup>. Il cosiddetto "Schema di Winograd" è costituito, quindi, da una domanda tripartita, così strutturata:

- una frase che contenga due sintagmi nominali e un pronome ambiguo, che potrebbe formalmente riferirsi a entrambi i sintagmi;
- una domanda sul sintagma a cui si riferisce il pronome;
- due risposte relative ai sintagmi nominali.

Apparentemente, si tratterebbe della semplice risoluzione di un'anafora; in realtà la conoscenza di fondo e il senso comune sono strumenti indispensabili all'individuazione della soluzione esatta.

«The trophy would not fit in the brown suitcase because it was too big. What was too big?

Answer 0: the trophy.

Answer 1: the suitcase».

Formalmente, il pronome neutro contenuto nella prima frase potrebbe riferirsi tanto a *trophy* quanto a *suitcase*; appare evidente, però, che, per via dei meccanismi che caratterizzano il funzionamento della realtà, solo una è la risposta plausibile: il pronome *it* non può che sostituire *trophy*, perché un trofeo troppo grande potrebbe non entrare nei volumi di una valigetta, laddove una valigetta troppo grande potrebbe contenerlo fin troppo.

«Since it was raining, I carried the newspaper over my backpack to keep it dry. What was I trying to keep dry?

Answer 0: the backpack.

Answer 1: the newspaper».

Anche in questo caso, il secondo pronome *it* potrebbe ipoteticamente riferirsi a entrambi i sintagmi nominali, ma l'evocazione della scena è dirimente: la soluzione non può che essere la prima.

Tuttavia, in entrambi i casi, senza l'accesso al senso comune e limitandosi alla superficie linguistica, le risposte 0 e 1 risultano egualmente valide: il riconoscimento dell'erroneità di una delle due risposte non dipende da questioni formali, bensì da parametri extralinguistici, di verosimiglianza, di rispondenza alle dinamiche che caratterizzano la realtà.

<sup>26</sup> Hector Levesque, *The Winograd Schema Challenge*, Toronto: University of Toronto, 2011.

## Il problema della comprensione

Il problema precedentemente sollevato in termini di senso comune consente di riflettere sulla questione più ampia della comprensione che, nonostante la diffusa fascinazione – di certo alimentata dal lessico opaco, personificante e ingannevole ampiamente adottato nel mondo dell'Intelligenza Artificiale –, non sembra rientrare nelle capacità della macchina.

La meccanicità dei dispositivi di IA, concepiti su basi informatizzate e computabili, è celata da un velo di mitizzazione, infittito da «mesmerici effetti antropomorfici»<sup>27</sup> che sono indizio del successo della simulazione. Ma l'identità dell'effetto non implica l'identità del sostrato che ne è la causa: il computer, anche quando dotato di una capacità di calcolo difficilmente concepibile per la mente umana, continua a lavorare in un modo rigorosamente computabile, molto diverso da quanto riconosciuto nella dinamica della mente dalle teorie più recenti; la complessità viene erroneamente scambiata per trascendenza.

Il computer «è solo volgare silicio [...] sono solo elettroni che vanno e vengono»<sup>28</sup>. Affermazioni simili sottintendono l'incommensurabilità tra uomo e macchina, una differenza ontologica giustificabile in vario modo. Limitatamente alla conoscenza attuale, l'impossibilità del confronto è motivata dalla diversa strutturazione della materia, che ha conseguenze sulla natura delle interazioni tra gli elementi: nel caso della macchina, l'interazione delle unità di materia sembrerebbe trovare piena spiegazione con la sola fisica classica, del tutto computabile; nel caso dell'uomo, né la fisica classica né la fisica quantistica tradizionale – tra l'altro, non computabile per definizione – sembrerebbero bastare e, secondo Roger Penrose, matematico britannico Nobel per la Fisica, bisognerebbe riferirsi a una fisica quantistica anomala per illustrare gran parte della fenomenologia della mente<sup>29</sup>.

Certamente sono innumerevoli i quesiti ancora aperti nell'ambito dello studio della mente e l'indagine scientifico-filosofica tenta di approssimarsi alle risposte attraverso l'adozione di prospettive teoriche diverse, sia in termini di multidisciplinarietà dei contributi, sia in termini di varietà delle concettualizzazioni della mente di riferimento – dagli approcci ormai minoritari, come quelli relativi a forme di dualismo o di riduzionismo estremo, fino al funzionalismo computazionale imperante, un fysicalismo non riduzionista<sup>30</sup>.

All'interno del dibattito, la questione della comprensione sarebbe focale in quanto componente funzionale della stessa intelligenza naturale: è il processo elaborativo, complesso e prolungato, attraverso cui razionalizziamo il mondo che ci circonda; è

<sup>27</sup> Henry Shevlin, *A Digital Remix of Humanity*, in: Justin Weinberg, *Philosophers On GPT-3 (updated with replies by GPT-3)*, «Daily Nous», 30 luglio 2020.

<sup>28</sup> Tito Orlandi, *Informatica testuale. Teoria e prassi*, cit., p. 164.

<sup>29</sup> Roger Penrose, *The Emperor's New Mind*, Oxford: Oxford University Press, 2016 [1989].

<sup>30</sup> Michele Di Francesco — Massimo Marraffa — Alfredo Tomasetta, *Filosofia della mente. Corpo, coscienza, pensiero*, Roma: Carocci, 2023 [2017].

lo strumento grazie al quale possiamo spostarci dal piano del presente, dell'analisi e dell'azione sincroniche, al piano del futuro, della pianificazione e della previsione; è il mezzo che permette di *generalizzare*, ossia di astrarre delle regolarità dietro le diverse apparenze fenomenologiche.

Sebbene le posizioni a riguardo siano molteplici, in generale la comprensione si delinea come un procedimento olistico che non emerge dal rispetto di regole algoritmiche.

Secondo Tito Orlandi – in analogia con la caratterizzazione dell'inconscio fornita dagli studi di scienza cognitiva<sup>31</sup> e in accordo con la concezione fodoriana della *sopravvenienza* del mentale sul fisico<sup>32</sup> – la comprensione sarebbe il risultato della combinazione di due tipologie di conoscenza, ambedue radicate nel soggetto che conosce, non nella cosa conosciuta:

- la conoscenza noetica, che è intrapsichica, non strutturata e, di conseguenza, non comunicabile;
- la conoscenza comunicabile, strutturata, ossia l'informazione<sup>33</sup>.

Dalle due definizioni di Orlandi si inferisce che l'unica conoscenza ontologicamente accessibile al computer è la seconda; una simile restrizione fa sì che l'informazione cui può accedere un calcolatore risulti essere parziale, incompleta rispetto a quanto decodificato dalla mente umana. Se solo la combinazione di una conoscenza formalizzabile, e quindi comunicabile, e una conoscenza intuitiva, non formalizzabile e quindi non comunicabile, può risultare nella comprensione, allora questa non può che restare preclusa alla macchina, per motivi costitutivi, poiché la macchina non ha i mezzi per elaborarla.

Se anche nell'ambito dell'Intelligenza Artificiale si parla ambigualmente di *apprendimento*, è bene sottolineare che quello della macchina è di tipo meramente meccanico, procedurale, matematico: non si svincola in nessun modo dai meccanismi materiali di computazione e non implica in nessun modo la comprensione. È una strategia algoritmica di miglioramento dei dati numerici, in particolare di riduzione della "funzione di errore", il cui modulo è inversamente proporzionale alla qualità del risultato dell'elaborazione.

Per via della natura del suo funzionamento e indipendentemente dalla complessità e dalla numerosità dei calcoli che è in grado di eseguire, la macchina non accede alla comprensione; una simile limitazione ha delle dirette conseguenze nei termini della decodifica e della ricodifica testuali.

<sup>31</sup> Massimo Marraffa — Alfredo Paternoster, *Sentirsi esistere. Inconscio, coscienza, autocoscienza*, Roma: Editori Laterza, 2013, p. 21: «i processi subpersonali della scienza cognitiva hanno caratteri diversi dalla coscienza: dove questa si presenta come unitaria, seriale, linguistica e sensibile a proprietà globali, quelli sono molteplici, paralleli, non linguistici e orientati all'elaborazione di proprietà locali».

<sup>32</sup> M. Di Francesco — M. Marraffa, A. Tomasetta, *Filosofia della mente. Corpo, coscienza, pensiero*, cit.

<sup>33</sup> T. Orlandi, *Informatica testuale. Teoria e prassi*, cit., p. 146-147.

In effetti, il computer è una mera macchina semiotica, un manipolatore di veicoli segnici: solo il soggetto umano, attraverso la sua competenza, caricandolo di significato, è in grado di trasformare il veicolo segnico in segno<sup>34</sup>. In altri termini, la conoscenza non è materialmente connessa ai dati codificati nella macchina, ma risiede nella mente, che reagisce attivamente allo stimolo del veicolo segnico, facendosi protagonista di un processo di decodifica-codifica. Per dirlo ancora con le parole di Tito Orlandi:

«il significato in senso proprio dei veicoli segnici gestiti dai computer, cioè i gruppi di bit, è attivo soltanto in presenza dei soggetti umani, mentre per i computer essi portano con sé un significato in senso molto particolare, cioè significano semplicemente se stessi, sono riconosciuti in quanto uno è diverso dagli altri [...]. La competenza umana rinvia invece al significato dei segni in quanto unità dell'informazione, diciamo così, esterna, cioè non ciò che essi sono, ma ciò che essi rappresentano»<sup>35</sup>.

Attraverso il glifo, l'occhio umano legge una rete di sensi, astrae le immagini e i contenuti evocati dal veicolo segnico materiale, che sta per qualcos'altro. Al contrario, per la macchina il glifo rappresenta solo sé stesso, un ente materiale che si combina con altri, innescando dei meccanismi di interazione e sostituzione con gli altri glifi come fossero pedine su una scacchiera.

La visione di Orlandi è coerente con quanto sostenuto da Searle in merito al rapporto tra sintassi e semantica e al confronto tra mente e Macchina di Turing: la semantica non è riducibile alla sintassi e dunque le operazioni mentali non possono essere definite in termini di manipolazioni meramente sintattiche<sup>36</sup>.

Dal punto di vista traduttologico, adottando questa cornice teorica di riferimento<sup>37</sup>, è evidente che la limitatezza della parte formalizzabile dell'apparato conoscitivo connesso al testo si riflette nell'incompletezza delle informazioni computabili, e dunque veicolabili, dalla macchina. Elementi di conoscenza noetica saranno sempre presenti, sebbene in tasso variabile e in modo più o meno evidente, anche in ragione della tipologia testuale in analisi: testi di natura diversa saranno, di conseguenza, più o meno adatti a essere sottoposti al trattamento automatico, e richiederanno un *editing* umano più o meno marcato, che possa colmare o comunque gestire al meglio le lacune informative. Ad esempio, il rapporto di biunivocità tra significante e significato che caratterizza gran parte del lessico utilizzato in un testo specialistico o in un testo pratico non è generalmente rintracciabile in un testo

<sup>34</sup> Ivi, p. 152.

<sup>35</sup> Ivi, p. 148.

<sup>36</sup> John R. Searle, voce *Chinese Room Argument*, in Robert A. Wilson — Frank C. Keil, *The MIT Encyclopedia of the Cognitive Sciences* (MITECS), Cambridge, Mass.: MIT Press, 1999, p. 115-116.

<sup>37</sup> Per approfondimenti: Dino Buzzetti, *Turing e l'elaborazione automatica del testo*, in: Accademia Nazionale dei Lincei, *Per il centenario di Alan Turing fondatore dell'informatica*, Roma: Scienze e Lettere, 2014, p. 159-188.

letterario o in un testo giornalistico, tipologie testuali che sfruttano l'ambiguità, in cui si ricorre ampiamente a giochi sulla polisemia o sulla forma delle parole.

### Una strada sbagliata?

Se le teorie sulle anomalie quantomeccaniche della mente di Penrose e sull'irriducibilità della semantica alla sintassi di Searle fossero corrette, allora le neuroscienze e il settore dell'Intelligenza Artificiale sarebbero attualmente in un vicolo cieco: nella prospettiva della modellizzazione dell'intelligenza generale, la riproduzione del connettoma biologico mediante le reti neurali sarebbe insufficiente, perché ne tralascerebbe una parte costitutiva essenziale; ma l'inclusione di tale componente fondamentale richiederebbe un kuhniano cambiamento di paradigma, che consentisse il superamento dei limiti imposti dalla computabilità<sup>38</sup>.

Focalizzando l'attenzione sulle basi ideologiche rintracciabili nell'approccio attuale allo studio di quanto concerne la mente umana e alla progettazione dei dispositivi di IA, sembra plausibile ricondurre l'insufficienza dei risultati finora conseguiti a due *bias* fondamentali, due preconcetti latenti fuorvianti che andrebbero problematizzati e contestati:

- che l'intelligenza umana sia del tutto, o quantomeno in massima parte, inquadabile in termini di razionalità, secondo la «secolare e imperturbabile formula dell'uomo "animale razionale"»<sup>39</sup>;
- che l'intelligenza umana rappresenti una forma di pensiero gerarchicamente superiore alle altre possibili.

A questo proposito, alcuni esponenti della comunità scientifica si mostrano scettici riguardo l'utilità dell'approccio imitativo in generale. Secondo Rodney Brooks, professore emerito di robotica presso il Massachusetts Institute of Technology (MIT), per esempio, il limite delle neuroscienze e del settore IA, ancor prima che pratico – il che sarebbe tanto più vero nel caso in cui la teoria di Penrose non venisse falsificata –, sarebbe concettuale e in qualche modo autoimposto: gli intenti imitativi adottati da gran parte della ricerca sarebbero controproducenti, perché, oltre a essere un riflesso del narcisismo umano, ingabbierebbero la mente dei ricercatori, adeguandola a uno schema di pensiero prestabilito. In questo modo sarebbe loro preclusa la libertà di concepire dei modelli di razionalità del tutto nuovi, svincolati dai meccanismi della tanto venerata mente umana. «*We are in an intellectual cul-de-sac, in which we model brains and computers on each other, and so prevent ourselves from having deep insights that would come with new models*»<sup>40</sup>.

<sup>38</sup> Per approfondimenti: Will Knight, *One of the fathers of AI is worried about its future*, «MIT Technology Review», 17 novembre 2018; Gary Marcus, *The deepest problem with deep learning*, «Medium», 1 dicembre 2018.

<sup>39</sup> Vallori Rasini, *L'essere umano. Percorsi di antropologia filosofica contemporanea*, Roma: Carocci, 2008, p. 120.

<sup>40</sup> Rodney Brooks, *Avoid the cerebral blind alley*, in: *Turing centenary. Is the brain a good model for machine intelligence?* «Nature» (22 febbraio 2012), n. 482, p. 462-463, <<https://doi.org/10.1038/482462a>>.

Come riporta Johannes Bruder, direttore dell'Istituto di ricerca in Design Sperimentale e Media Culture (IXDM) di Basilea:

«While the analogy between human brain and computer might have been fruitful for both computing and neuroscience over the course of the 20th century, the belief that computers could become “intelligent” could now hold artificial intelligence back»<sup>41</sup>.

Che i propositi imitativi, anche alquanto narcisistici, costituiscano un impedimento e limitino davvero i risultati, pratici e teorici, ottenibili, soffocando la creatività necessaria al progresso?

*This article contributes to heated debate around the multidisciplinary field of Artificial Intelligence by adopting a specific perspective: the translational point of view. Machine Translation was the first application of computer technology in the humanities, both because of the general interest in the study of human language, given its strong identity connotation, and because of its high formalization. Through an overview of the architectures that have followed, the paper focuses on the fruitful dialogue between practice and theory, which triggered modern reflections on the mechanisms of translation as a decision-making and negotiation process. These include: the limits of the interlingual transfer, the entropy of information, the anisomorphism of natural languages, the mutual influence of language and cognition, the relevance of extratextual information and speakers' cognitive horizons, the concept of textual fidelity. The deepening of these aspects opens up wider reflections, such as the thorny question of the relationship between natural and artificial intelligence, the essence of understanding, and the possible heuristic impediment constituted by imitation.*

<sup>41</sup> Johannes Bruder, *Infrastructural intelligence: Contemporary entanglements between neuroscience and AI*, «Progress in Brain Research», 233 (31 luglio 2017), cap. 5, p. 101-128, <<https://doi.org/10.1016/bs.pbr.2017.06.004>>.

L'ultima consultazione dei siti web è avvenuta nel mese di giugno 2024