

Recensione: Claudio Paolucci, *Nati cyborg. Cosa l'intelligenza artificiale generativa ci dice dell'essere umano*. Roma: Luca Sossella editore, 2025.

Fabio Ciotti

Università di Roma - Tor Vergata, Italia
fabio.ciotti@uniroma2.it

Il dibattito sull'intelligenza artificiale generativa ha prodotto, in un arco di tempo relativamente breve, una letteratura molto ampia, distribuita tra ricostruzioni tecniche, analisi delle possibili trasformazioni sociali, politiche ed economiche, riflessioni sugli impatti etici e sulla natura dei modelli linguistici dal punto di vista filosofico. Il livello teorico di questa produzione è molto diseguale e, d'altronde, la polarizzazione del dibattito scientifico e pubblico sull'intelligenza artificiale non favorisce quella profondità di riflessione che il fenomeno richiederebbe. Merita, perciò, attenzione un volume volutamente agile nelle dimensioni e nello stile comunicativo, ma denso nei contenuti: *Nati cyborg. Cosa l'intelligenza artificiale generativa ci dice dell'essere umano*, di Claudio Paolucci, filosofo del linguaggio e semiologo, allievo di Eco, che negli ultimi anni ha esplorato una delle declinazioni teoricamente più produttive della semiotica, la semiotica cognitiva.

Nella parte teoricamente più rilevante del dibattito sugli LLM si distinguono due correnti: chi cerca di collocare la comprensione dell'IA entro il quadro di una teoria della mente e degli agenti cognitivi, di tradizione analitica, e chi invece ragiona sugli LLM come macchine del linguaggio, rifacendosi alla tradizione della linguistica teorica che da Saussure arriva alla teoria semiotica e alle varie correnti del poststrutturalismo. Il libro di Paolucci si iscrive in questo secondo orizzonte teorico, in cui si possono collocare anche il volume di Leif Weatherby *Language Machines: Cultural AI and the End of Remainder Humanism* (Weatherby 2025), che legge gli LLM come convergenza di computazione e linguaggio e li interpreta non come simulatori di cognizione ma come produttori di cultura, denunciando come «remainder humanism» la difesa di un'essenza umana definita per sottrazione rispetto alla tecnica, e il libro di N. Katherine Hayles *Bacteria to AI: Human Futures with Our Nonhuman Symbionts* (Hayles 2025), che propone un «Integrated Cognitive Framework» in cui la cognizione è ridefinita in senso relazionale, come processo che interpreta informazione in contesti dotati di significato, e viene estesa lungo un continuum che va dai batteri all'intelligenza artificiale, con un intento dichiaratamente anti-anthropocentrico. Rispetto a questi lavori, *Nati cyborg* condivide il presupposto di fondo, il rifiuto di una linea di demarcazione essenzialista fra umano e macchinico, ma vi giunge per una via propria, quella della semiotica dell'enunciazione e della semiotica cognitiva, e si distingue per la

scelta di assumere l'intelligenza artificiale non soltanto come oggetto di indagine, ma come dispositivo teorico attraverso cui riesaminare alcune categorie tradizionalmente impiegate per descrivere l'essere umano: linguaggio, significato, intelligenza, soggettività e azione.

1. Dal cyborg alla teoria dell'enunciazione

La tesi centrale del libro è formulata fin dalle pagine iniziali: la trasformazione più rilevante prodotta dall'intelligenza artificiale generativa non riguarda solo e tanto le macchine, ma la rappresentazione che gli esseri umani hanno di se stessi. La comparsa sulla scena della comunicazione sociale di sistemi non biologici capaci di produrre enunciati complessi mette in crisi l'identificazione esclusiva tra linguaggio e natura umana. L'intelligenza artificiale generativa è una svolta che, come suggerisce l'epigrafe nietzschiana posta in apertura, rivela come l'essere umano sia da sempre un cyborg: un essere costitutivamente debole, *in-baculum* nell'etimologia di «imbecille» suggerita da Ferraris (Ferraris 2016), che si emancipa dalla propria fragilità delegando all'ambiente, agli utensili, alle scritture, alle tecnologie, porzioni crescenti del proprio lavoro cognitivo. D'altra parte, il titolo del libro allude esplicitamente alle teorie di Andy Clark e al suo ben noto *Natural-Born Cyborgs* (Clark 2003), di cui il libro si dichiara un'estensione alla luce dell'IA generativa. La scrittura, gli archivi, le tecniche di calcolo, le istituzioni e i dispositivi digitali non rappresentano semplici aggiunte a una mente già precostituita. Sono componenti dei sistemi attraverso cui gli esseri umani ricordano, inferiscono, progettano e agiscono.

L'impianto anti-antropocentrico di Paolucci procede smontando l'assunto che ricorre in gran parte della letteratura critica sull'IA: l'assunto secondo cui ciò che le macchine fanno non sarebbe vero pensiero, vero linguaggio, vera comprensione, perché difetterebbe di coscienza, intenzionalità e significato nel modo genuinamente umano. A questo assunto il libro oppone un argomento che ricorre in molte varianti lungo tutto il testo: è necessario smettere di porre il modo umano di fare le cose a campione-parametro rispetto al quale misurare, e di norma squalificare, tutto ciò che da esso differisce. Le macchine intelligenti operano diversamente, senza dubbio, e possono fare meglio o peggio dell'uomo, ma il punto veramente interessante è capire cosa il loro modo di farle riveli del nostro.

La teoria dell'enunciazione costituisce il primo strumento con cui Paolucci affronta il problema. In Benveniste l'enunciazione è l'atto attraverso il quale un parlante si appropria della lingua e si istituisce come soggetto, assumendo la posizione dell'«io» (Benveniste 1971). Paolucci riprende questo quadro, ma contesta che l'agency dell'atto linguistico possa essere ricondotta integralmente a un soggetto individuale. La teoria impersonale dell'enunciazione, già elaborata nel suo precedente e ben più ponderoso volume, *Persona*, descrive ogni enunciato come il risultato dell'assemblaggio di una pluralità di istanze: norme linguistiche e sociali, generi discorsivi, istituzioni, abitudini, stereotipi, formule ed enunciati precedenti (Paolucci 2020). Il parlante concatena la propria voce a tali istanze, ma non ne costituisce l'unica origine. L'agency enunciativa è quindi distribuita, senza che ciò implichi l'eliminazione del parlante umano. Questa precisazione è rilevante per comprendere l'estensione della teoria agli LLM. Paolucci non sostiene semplicemente che non esista alcun soggetto dietro gli enunciati umani. Sostiene piuttosto che la presenza di un soggetto individuale non esaurisce le condizioni di produzione del discorso. Se anche l'enunciazione umana dipende dall'attualizzazione e dalla ricombinazione di materiali linguistici e culturali preesistenti, la dipendenza degli LLM dai dati di addestramento non basta a escludere i loro prodotti dal dominio dell'enunciazione.

La conseguenza è una ridefinizione della domanda «chi parla?». Nel caso di una macchina generativa, l'assenza di una persona biologica non comporta l'assenza di un'istanza enunciante. La macchina produce enunciati a partire da una struttura nella quale intervengono il modello, i dati, le procedure di allineamento, il prompt, l'utente e il contesto d'uso. L'enunciazione macchinica non coincide pertanto con l'enunciazione umana, ma può essere descritta come una diversa configurazione della medesima funzione semiotica.

2. Il significato e l'enciclopedia

La parte teoricamente più densa del libro è quella in cui l'influenza di Umberto Eco si fa più evidente: la critica di quello che Paolucci chiama il «mito del significato». L'argomento prende le mosse da un'interpretazione semiotica del test di Turing (Turing 1950) per poi affrontare il controargomento classico contro l'intelligenza delle macchine, il *Gedankenexperiment* della stanza cinese di Searle: la macchina manipolerebbe soltanto significanti sul piano sintattico, secondo regole formali, senza mai accedere alla semantica, che resta prerogativa della mente umana (Searle 1980). Paolucci tratta la stanza cinese come l'argomento ancora largamente insuperato di chi nega all'IA ogni reale capacità linguistica, e ne riconosce la forza, ma la risposta che oppone non consiste nel sostenere che la macchina comprende come l'uomo. Consiste, più radicalmente, nel mettere in discussione la nozione stessa di significato che l'argomento dà per scontata, e con essa alcune distinzioni, come quella tra significante e significato, o tra sintassi e semantica, che ne sono tanto il presupposto quanto la conseguenza. Contro queste opposizioni Paolucci, come già Eco a suo tempo, mobilita in primo luogo il pragmatismo di Peirce, secondo cui il contenuto di un concetto deve essere ricondotto alle sue conseguenze pratiche concepibili e alle abitudini d'azione che esso organizza (Peirce 1931–1958, 5.402). Da questa prospettiva, significare non è il possesso di un'entità mentale interna, ma una funzione che si manifesta negli usi, nelle inferenze e negli effetti di un segno.

Il capitolo passa poi a mettere al lavoro la teoria della semantica come enciclopedia di Eco, nella sua lunga evoluzione. Nel *Trattato di semiotica generale* (Eco 1975), riformulando il modello semantico di Katz e Fodor (Katz e Fodor 1963), Eco proponeva di sostituire alla struttura componenziale a dizionario, chiusa, un'enciclopedia virtualmente illimitata, nella quale il significato di un termine dipende dall'attivazione contestuale di percorsi e selezioni differenti. Paolucci sostiene che il Modello Semantico Riformulato descriva in modo particolarmente efficace alcune operazioni dei modelli linguistici contemporanei. Un LLM, come noto, non conserva una copia letterale dei testi di addestramento, ma una rappresentazione distribuita e compressa delle regolarità linguistiche e culturali presenti nei dati. Di fronte a un prompt, il sistema attualizza alcune possibilità e ne inibisce altre, selezionando una continuazione in funzione del contesto. Il significato viene così descritto come un'operazione di riduzione della complessità. Ogni sistema finito deve selezionare, comprimere e organizzare l'informazione disponibile. La metafora del «JPEG sfocato», ripresa da Ted Chiang, indica precisamente il carattere della rappresentazione prodotta dai modelli generativi (Chiang 2023) e Paolucci generalizza tale intuizione: anche la percezione e la semantica umane dipendono da processi selettivi che non riproducono integralmente il mondo, ma ne estraggono configurazioni pertinenti per l'azione.

L'analisi di Paolucci evita di assumere il significato umano come unico modello possibile di semantica, e non richiede che il processo umano e quello macchinico siano identici. In termini di capacità osservabili, senza dubbio, gli LLM mostrano prestazioni sistematiche di parafrasi, disambiguazione, adattamento contestuale, inferenza e manipolazione controfattuale. Tali capacità costituiscono evidenze pertinenti per l'attribuzione di una competenza semantica, senza

che sia necessario identificarla con l'esperienza umana del significato. Resta ovviamente il problema teorico di stabilire se la comune capacità di selezionare, disambiguare e riorganizzare un'enciclopedia compressa sia sufficiente, sul piano funzionale, per attribuire in modo sensato ai due sistemi una proprietà semantica dello stesso genere. Oppure se, inversamente, lo stesso Modello Semantico Riformulato sia un modello che ipostatizza in modo astratto un più profondo modo di funzionare di un agente semantico, in cui la nozione stessa di rappresentazione, pur operante in tale modello, diventa una categoria derivata e non primaria, un modo fin troppo umano per rendere conto di un sistema che opera a basso livello secondo principi affatto differenti, che le reti neurali probabilistiche degli LLM catturano in modo più efficace.

3. L'«intelligenza zero» di Floridi

Lasciando queste suggestioni alla riflessione dei lettori, conviene spendere alcune parole su quello che è probabilmente il principale confronto polemico del libro di Paolucci, quello contro la nozione di «intelligenza zero» formulata da Luciano Floridi nel suo recente *La differenza fondamentale*, che riprende e sviluppa le tesi di un saggio del 2023 (Floridi 2025). Nella ricostruzione di Paolucci, la tesi di Floridi si fonda su un decoupling tra agency e intelligenza: un sistema artificiale può svolgere efficacemente un compito senza comprendere ciò che fa e senza essere intelligente. La riuscita della macchina dimostrerebbe quindi che il compito poteva essere realizzato anche «a intelligenza zero».

Paolucci contesta questa posizione su tre piani. In primo luogo, nega che l'efficacia dell'azione possa essere separata in modo netto dall'intelligenza. Il pragmatismo e l'enattivismo descrivono infatti il pensiero come una capacità orientata alla soluzione di problemi e alla regolazione dell'azione. In secondo luogo, individua nel concetto di intelligenza zero un presupposto antropocentrico: il funzionamento umano viene assunto come parametro e ogni deviazione viene interpretata come mancanza. In terzo luogo, osserva che, una volta separata l'intelligenza dall'azione efficace, diventa difficile stabilire positivamente che cosa il termine «intelligenza» designi. A queste posizioni oppone una versione del principio di parità di Clark: se un processo, svolto da un pezzo di mondo, produce effetti che giudicheremmo intelligenti qualora fossero prodotti da un umano, esso va trattato come intelligente. Si può convenire senza residui con queste critiche, poiché anche a chi scrive l'intero apparato teorico di Floridi appare un investimento concettuale in larga parte mal riposto, che impiega un considerevole lavoro filosofico per sostenere una distinzione, quella fra agency e intelligenza, la cui plausibilità dipende dalla riduzione della nozione di agency a quella di azione, oppure da una non esplicitata adesione alla Actor-Network Theory di Latour (Latour 2005) o persino all'ontologia orientata agli oggetti di Harman (Harman 2018).

La critica di Paolucci alle tesi di Floridi individua, peraltro, una strategia comune nelle posizioni deflazionistiche applicate all'intelligenza artificiale. Se ogni compito precedentemente considerato intelligente viene retroattivamente ridefinito come non intelligente nel momento in cui una macchina riesce a svolgerlo, la tesi dell'assenza di intelligenza diventa difficilmente falsificabile. La separazione tra azione efficace e intelligenza rischia allora di operare come premessa piuttosto che come conclusione, e significa decidere in anticipo, e non scoprire, ciò che si pretende di dimostrare: che la macchina non pensa. Da ciò non segue che ogni comportamento efficace debba essere considerato intelligente. È necessario distinguere tra un meccanismo causalmente efficace e un sistema capace di adattarsi a contesti variabili, integrare informazioni, generalizzare e modificare la propria condotta in relazione a obiettivi e vincoli. La

questione deve essere risolta attraverso l'analisi delle capacità e dell'organizzazione del sistema, non attraverso la sua classificazione preliminare come biologico o artificiale.

Il caso di AlphaGo svolge nel libro una funzione paradigmatica in questo senso (Silver et al. 2017). La rilevanza filosofica della vittoria nel Go non consiste soltanto nel superamento della prestazione umana, ma nella produzione di mosse inizialmente giudicate prive di senso dagli esperti e successivamente riconosciute come strategicamente efficaci. Il sistema rende visibili possibilità che la tradizione del gioco non aveva esplorato e modifica, di conseguenza, il modo in cui gli esseri umani comprendono il gioco stesso.

4. Un libro che vale la pena leggere

In conclusione, *Nati cyborg*, fra i molti libri recenti dedicati all'IA generativa, è uno dei pochi che uniscano densità teorica, efficacia argomentativa e una commendabile brevità, e quasi l'unico, nel panorama italiano, a far valere con serietà il punto di vista di una tradizione di pensiero che ha avuto in passato una grande influenza, ma che negli ultimi anni è stata troppo facilmente dimenticata. Il suo valore principale non sta tanto nelle singole tesi, alcune delle quali possono essere esposte a obiezioni, quanto nel gesto metodologico complessivo: rovesciare la domanda, smettere di chiedere se la macchina sia all'altezza dell'uomo e cominciare a chiedere cosa la macchina riveli dell'uomo.

È su questo gesto che conviene chiudere, anche perché è il punto in cui il libro chiama in causa il lettore che si occupi degli stessi problemi. Rispetto all'impianto di Paolucci si può infatti assumere una posizione diversa nei presupposti teorici e tuttavia convergente nella diagnosi. La decostruzione dell'io che il libro conduce per via linguistica può essere condotta anche per altra strada, quella di alcuni filosofi che appartengono alla tradizione della filosofia analitica del linguaggio e della mente: dalla critica wittgensteiniana del mentalismo interiore, con l'argomento contro il linguaggio privato e la concezione del significato come uso pubblico (Wittgenstein 2009), e dalla dissoluzione dennettiana dei misteri dell'intenzionalità e dell'esperienza soggettiva in favore dell'atteggiamento intenzionale (Dennett 1987) e della critica concettuale ai miti del teatro cartesiano e dei qualia. Sono due strategie di pensiero diverse: la tradizione linguistica e semiotica, mettendo tra parentesi la natura mentale dei fatti relativi al linguaggio, rischia di espungere anche le domande legittime sulla natura dei fatti mentali che non si lasciano ridurre all'enunciazione; a via wittgensteiniana e quella dennettiana (che sono tutt'altro che facilmente sovrapponibili, sia chiaro) nega che i fenomeni mentali e linguistici abbiano una natura metafisicamente interiore e sostanziale senza per questo negare che vi sia una mente di cui dare conto. Ma è significativo che, muovendo da Wittgenstein e Dennett da un lato e da Peirce ed Eco dall'altro, con Clark a fare da mediatore, si pervenga a conclusioni largamente convergenti: che nulla di nascosto e irriducibile si dia da scoprire dietro i comportamenti dei sistemi intelligenti, naturali o artificiali che siano, e che i criteri in base ai quali si attribuiscono competenze a un agente artificiale siano gli stessi in base ai quali le si attribuiscono, o si negano, a chiunque. La struttura linguistica, in questa prospettiva, è ciò che con Dennett si potrebbe chiamare un *real pattern* (Dennett 1991), un'astrazione di livello superiore scientificamente legittima, e ciò vale per i sistemi biologici come per quelli artificiali.

Ciò che in ultima analisi conta, ed è il merito di *Nati cyborg*, non è la scelta dell'una o dell'altra cornice teorica, ma il superamento dell'antropocentrismo e la costruzione di una teoria della cognizione, del linguaggio e del significato ontologicamente più inclusiva: una teoria capace di rendere conto, con le medesime categorie, dei processi semiotici e cognitivi che attraversano i sistemi biologici e quelli artificiali, senza erigere quel poco che sappiamo della natura umana a

misura esclusiva del pensiero e del senso, poiché ormai siamo in un mondo abitato anche da agenti non umani.

Riferimenti bibliografici

- Benveniste, Émile. 1971. *Problemi di linguistica generale*. Traduzione di Maria Vittoria Giuliani. Milano: Il Saggiatore.
- Chiang, Ted. 2023. «ChatGPT Is a Blurry JPEG of the Web». *The New Yorker*, 9 febbraio 2023.
- Clark, Andy. 2003. *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford: Oxford University Press.
- Dennett, Daniel C. 1987. *The Intentional Stance*. Cambridge, MA: MIT Press.
- Dennett, Daniel C. 1991. «Real Patterns». *The Journal of Philosophy* 88 (1): 27–51.
- Eco, Umberto. 1975. *Trattato di semiotica generale*. Milano: Bompiani.
- Ferraris, Maurizio. 2016. *L'imbecillità è una cosa seria*. Bologna: Il Mulino.
- Floridi, Luciano. 2025. *La differenza fondamentale. Artificial Agency: una nuova filosofia dell'intelligenza artificiale*. Milano: Mondadori.
- Harman, Graham. 2018. *Object-Oriented Ontology: A New Theory of Everything*. London: Pelican Books.
- Hayles, N. Katherine. 2025. *Bacteria to AI: Human Futures with Our Nonhuman Symbionts*. Chicago: University of Chicago Press.
- Katz, Jerrold J., e Jerry A. Fodor. 1963. «The Structure of a Semantic Theory». *Language* 39 (2): 170–210.
- Latour, Bruno. 2005. *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford University Press.
- Paolucci, Claudio. 2020. *Persona: Soggettività nel linguaggio e semiotica dell'enunciazione*. Milano: Bompiani.
- Paolucci, Claudio. 2025. *Nati cyborg: Cosa l'intelligenza artificiale generativa ci dice dell'essere umano*. Roma: Luca Sossella editore.
- Peirce, Charles S. 1931–1958. *Collected Papers of Charles Sanders Peirce*. Voll. 1–6 a cura di Charles Hartshorne e Paul Weiss; voll. 7–8 a cura di Arthur W. Burks. Cambridge, MA: Harvard University Press.
- Searle, John R. 1980. «Minds, Brains, and Programs». *Behavioral and Brain Sciences* 3 (3): 417–457.
- Silver, David, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, et al. 2017. «Mastering the Game of Go without Human Knowledge». *Nature* 550 (7676): 354–359.
- Turing, Alan M. 1950. «Computing Machinery and Intelligence». *Mind* 59 (236): 433–460.
- Weatherby, Leif. 2025. *Language Machines: Cultural AI and the End of Remainder Humanism*. Minneapolis: University of Minnesota Press.

Wittgenstein, Ludwig. 2009. *Philosophical Investigations*. 4^a ed. A cura di P. M. S. Hacker e Joachim Schulte. Traduzione di G. E. M. Anscombe, P. M. S. Hacker e Joachim Schulte. Chichester: Wiley-Blackwell.