Università degli Studi Suor Orsola Benincasa

ANNALI

2024



ANNALI

2024

SUOR ORSOLA
BENINCASA
UNIVERSITÀ EDITRICE

Direttore
Vincenzo Omaggio

Direttore responsabile
Arturo Lando

Comitato scientifico
Lucio d'Alessandro
Francesco M. De Sanctis
Enricomaria Corbi
Tommaso Edoardo Frosini
Emma Giammattei
Aldo Sandulli
Paola Villani

Redazione
Luciana Trama (responsabile)
Flavia Soprani

Gli articoli negli «Annali» dell'Università Suor Orsola Benincasa sono preliminarmente sottoposti a un giudizio di referaggio nel rispetto dell'anonimato sia dell'autore che dei revisori (a doppio cieco). La valutazione è affidata a uno o più specialisti del tema trattato, designati dal Direttore degli «Annali» anche fra i componenti del Comitato Scientifico, che rivestano, o abbiano rivestito, la posizione di professore universitario nelle Università italiane o straniere ovvero si distinguano come esperti di chiara fama.

I revisori giudicano, entro 30 giorni dal ricevimento dell'elaborato, sull'opportunità di accogliere il lavoro scientifico tenendo conto dei parametri valutativi di cui alla vigente legislazione e formulando un giudizio motivato – senza modifiche o previo apporto di modifiche – di accoglimento o di rigetto. Il Direttore, infine, su propria responsabilità, può decidere di non assoggettare a revisione i lavori di particolare pregio o di autori di particolare prestigio.

—

Progetto grafico e impaginazione
Flavia Soprani

Realizzazione e stampa
Carmine Marra

© Università degli Studi Suor Orsola Benincasa, Napoli 2024
Tutti i diritti sono riservati

ISSN *online* 2281-3241
ISSN *press* 2037-5867

SOMMARIO

7 GENNARO CATONE MIRELLA DONATO
Cancer-Related Fatigue Syndrome *(CRFS): il dolore del corpo e la sofferenza dell'anima*

45 CLAUDIA CERNICCHIARA GIOVANNI COPPOLA PIERO SCOGNAMIGLIO
Sul graffito perduto del labirinto della casa di Marco Lucrezio a Pompei

69 EMANUELE COLUCCIA
L'apprendimento di percorsi

109 EDOARDO D'ANGELO
La Basilicata nella documentazione fridericiana

121 ALMERINDA DI GAETANO
Il contributo della linguistica educativa alla formazione professionale: uno studio pilota

137 RICCARDO FINOCCHI ANTONIO PERRI
Le menti (e i programmi) significano? Forme di vita e Artificial Intelligence

155 ANTONELLA GRITTI TIZIANA SALVATI MARTINA ERRICO GENNARO CATONE
Consultazioni con studenti universitari: un osservatorio sulle forme di disagio giovanile

175 ANTONIO LUONGO
Capograssi e l'esperienza giuridica', ovvero dalla "preistoria spirituale" alla positivizzazione del diritto

193 DANIELA MANNO DIANA CATALDO
Diventare autori dell'inclusione. Appunti per la formazione degli insegnanti

217 SARA NAVATTI
L'uscita dall'anonimato: testimonianze di firme nella ceramica e nella statuaria dell'età greca arcaica

263 MARCO NIELI
Pigmei, Numidi, "Etiopi": per una lettura del tema "africano" nella produzione di IV stile a Pompei

303 GIUSEPPE PERTA
Napoli e il suo golfo nei portolani tardomedievali

325 PASQUALE ROSSI
Formazione professionale e tutela del territorio. Conservazione e restauro dei beni culturali all'Università Suor Orsola Benincasa: il valore della terza missione per la città di Napoli e dei suoi dintorni

345 NATASCIA VILLANI
Alle origini dell'etica della cura

357 ANTONIO ZAPELLI
De merito. *Per la critica di un'ideologia e delle sue retoriche*

RICCARDO FINOCCHI ANTONIO PERRI *

Le menti (e i programmi) significano? Forme di vita e Artificial Intelligence

L'articolo analizza alcune pratiche e forme di vita emerse nell'ambito dell'AI e dei dispositivi digitali, a partire da riflessioni delle *Ricerche filosofiche*. Queste ultime verranno ripensate in relazione all'ormai pluridecennale dibattito sull'"intelligenza" dei computer, prefigurato da Turing e ulteriormente sviluppato nel 1980 dall'esperimento mentale della *Chinese room* di Searle.

Parole chiave: Assistenti virtuali, Intelligenza artificiale, Test di Turing, giochi linguistici, forme di vita

The aim of this paper is to explore some practices and Lebensformen *which have emerged in interactions with AI-driven digital devices, moving from Wittgenstein's seminal thoughts about peculiar language-games (*Sprachespielen*) in his* Philosophische Untersuchungen. *Pre-digital insights of the Austrian philosopher will be rephrased in the light of the controversial debate about computer's "intelligence" – which Turing foreshadowed, and Searle further developed during the Eighty with his famous thought experiment of the Chinese Room.*

Keywords: Virtual assistants, Artificial Intelligence, Turing test, language games, forms of life

* Gli autori hanno discusso assieme i contenuti di questo contributo. Per ciò che riguarda la stesura materiale dei paragrafi, a Riccardo Finocchi si devono i §§ 1 e 3 mentre Antonio Perri ha scritto i §§ 2 e 4. Le *Conclusioni* sono state redatte congiuntamente.

1. *Introduzione: menti artificiali e forme di vita nell'epoca dell'AI*

La nostra quotidianità è popolata di oggetti digitali basati su Intelligenza Artificiale (*Artificial Intelligence* – AI) che 'dialogano' con i loro utenti (umani e non) attraverso segnali manifestati da sostanze espressive diverse: l'interazione verbale (*Natural Language Processing*) con gli *home virtual assistants* o le *self driving cars*; la visione in rete di *deepfakes*, immagini di sintesi prodotte dall'intelligenza artificiale; le informazioni veicolate dagli algoritmi nei *social media*; *virtual reality* e *augmented reality* che si basano sempre di più sulle potenzialità dell'intelligenza artificiale; suoni elaborati da programmi musicali *AI-driven*; fino alla processazione di veri e propri testi scritti o, meglio, generati da algoritmi complessi basati sui cosiddetti *Large Language Models* come *ChatGPT* – in grado, come noto, di scrivere tesi o articoli scientifici[1]. Si

[1] Dal punto di vista tecnico, ovviamente, non siamo in grado di entrare nei dettagli relativi al funzionamento delle singole tecnologie evocate, che in effetti sono diversi e non tutti assimilabili a un'unica logica algoritmica. In particolare, i *virtual* o *vocal assistants*, sino a tempi recenti, non si giovavano di accesso a grandi quantità di dati da processare per fornire risposte a interrogazioni da parte dell'utente umano (ma, nondimeno, costituiscono ancora un ramo importante dell'NLP, configurando il settore denominato *speech recognition* e *speech processing*). Il funzionamento di Siri o Alexa è stato sinora in larga misura simile a quello di un qualunque motore di ricerca che recuperi informazioni indicizzate dalla Rete, e non è dunque assimilabile a quello fondato sul *machine learning* e algoritmi complessi caratteristico dei c.d. *Large Language Models (LLM)* e di agenti conversazionali come *ChatGPT*. Riteniamo tuttavia che queste differenze si vadano progressivamente assottigliando, e in prospettiva si annulleranno a breve. Non a caso il «New York Times» del 10 giugno 2024 dava notizia di un nuovo progetto di Apple: usare l'AI generativa di *OpenAI*, che ha creato *ChatGPT*, per dar vita a ciò che l'azienda già chiama *Apple Intelligence*; si tratta di una decisione che, secondo il NYT, «will also result in a major upgrade for Siri, the voice assistant that has languished» (*https://www.nytimes.com/2024/06/10/technology/apple-intelligence-openai.html*).

tratta di nuovi dispositivi (anche in senso focaultiano[2], posto che decidono del poter-fare di molti attori), che configurano universi di ibridazioni[3]) organizzando (e venendo organizzati in) forme di vita.

Il punto centrale di questo contributo, dunque, è tracciare una lettura dei processi, ormai pervasivi nelle pratiche quotidiane, dell'interazione uomo-AI attraverso una lettura wittgesteiniana, nella convinzione che questa sia una delle migliori strategie di comprensione del fenomeno in una prospettiva di filosofia del linguaggio. In sostanza, si tratta di applicare – lo dichiariamo fin da subito – all'analisi del rapporto uomo-AI alcune nozioni centrali delle *Ricerche filosofiche*[4], cercando di individuare un qualche tipo di *gioco linguistico* entro cui le reciproche attività 'intelligenti' (di uomo e macchina) *significano* (per gli attori umani). Come è noto, la nozione di giochi linguistici presuppone la dimensione relativa ai modi in cui il linguaggio viene usato, ovvero la nozione di significato come uso; infatti, come scrive De Mauro «le forme linguistiche hanno un significato perché sono usate dall'uomo, e solo in quest'uso trovano la garanzia di essere collegate a un determinato significato»[5]. Queste osservazio-

[2] Cfr. M. Foucault, *Le jeu de Michel Foucault* (1977), in Id., *Dits et Écrits* 1954-1988, t. III (1976-1979) Paris, Gallimard, 1994, pp. 298-328; G. Agamben, *Che cos'è un dispositivo*, Milano, Nottetempo, 2010.

[3] B. Latour, *Nous n'avons jamais été modernes*, Paris, La Découverte, 1991; trad. it. *Non siamo mai stati moderni*, Milano, Eleuthera, 1995; Id., *Where are the Missing Masses? The Sociology of a Few Mundane Artifacts*, in W. E. Bijker, J. Law (eds.), *Shaping Technology/Building Society: Studies in Sociotechnical Change*, Cambridge (Mass.), MIT Press, pp. 225-258; trad. it. *Dove sono le masse mancanti?*, in A. Mattozzi, a cura, *Il senso degli oggetti tecnici*, Roma, Meltemi, 2006, pp. 81-124; P. Peverini, *Inchiesta sulle reti di senso. Bruno Latour nella svolta semiotica*, Meltemi, Milano 2023.

[4] Cfr. L. Wittgenstein, *Philosophische Untersuchungen*, Oxford, Basil Blackwell, 1953; trad. it., *Ricerche filosofiche*, Torino, Einaudi, 1967 [nel testo citato come RF seguito dal n. del paragrafo].

[5] T. De Mauro, *Introduzione alla semantica*, Bari, Laterza, 1965, p. 184.

ni dovranno poi, giocoforza, essere rilette alla luce dell'irruzione dell'AI negli usi e pratiche linguistiche della contemporaneità, coinvolgendo al contempo la possibilità dello stabilirsi di regole come meccanismi di accordo tra parlanti (non necessariamente umani) entro forme di vita: «la parola gioco linguistico – scrive Wittgenstein (RF §23) – è destinata a mettere in evidenza il fatto che il parlare un linguaggio fa parte di un'attività, o di una forma di vita» – e ancora (RF §19) «immaginare un linguaggio significa immaginare una forma di vita».

Tempo addietro, tra il 1972 e il 1979, Hubert L. Dreyfus nel testo *What computer can't do*[6], aveva utilizzato anche temi wittgensteiniani per negare la possibilità di creare un *uomo artificiale*; era, questa, un'idea di AI forte in quegli anni oggetto delle ricerche condotte dal MIT, nel quale Noam Chomsky era a capo dell'area linguistica. Proprio l'impossibilità che la macchina potesse *comprendere l'uso* di enunciati entro giochi linguistici, oltre all'assenza di una dimensione 'corporea' o sensoriale delle macchine informatiche di quegli anni[7], erano sembrati a Dreyfus argomenti inoppugnabili. A esemplifi-

[6] H.L. Dreyfuss, *What Computers Can't Do: The Limits of Artificial Intelligence*, New York, Harper & Row, 1977 (rev. ed. 1979); trad. it. *Che cosa non possono fare i computer. I limiti dell'intelligenza artificiale*, Roma, Armando, 1988.

[7] Questa caratteristica è alla base del noto esperimento mentale di Putnam (H. Putnam, *Reason, Truth and History*, Cambridge, Mass., Cambridge University Press, 1981; trad. it. *Ragione, verità e storia*, Milano, Il Saggiatore, 1985) sui «cervelli in una vasca», nel quale il filosofo dimostrava l'impossibilità per un *device* sprovvisto di apparato sensoriale – come quello immaginato da Turing su cui cf. *infra* – di compiere atti di riferimento: siamo noi esseri umani, argomentava il filosofo, che abbiamo «una consuetudine di rappresentazioni per la quale il discorso del computer si riferisce alle mele, ai campanili, al paesaggio, e così via» (Putnam, cit., tr. it. 1985, p. 17). Occorre naturalmente chiedersi cosa cambi davvero nel momento in cui il *device* sia «provvisto di occhi e orecchie elettroniche», oltre che «dei mezzi per incorporare dati provenienti da tali organi sensoriali o per con-

care, tra l'altro, le sue affermazioni Dreyfus citava non a caso il gioco degli scacchi (esempio chiaramente *in auge* tra i filosofi del linguaggio, seguendo la linea che muove da Saussure e passa attraverso la teoria dei giochi): un *uomo artificiale* del MIT, sosteneva, è in grado di giocare a scacchi (quello dei giochi, in quegli anni, costituì un passaggio importante nello sviluppo dell'AI), perché, come osservato da Roncaglia, «indubbiamente una programmazione 'classica', strettamente logica e autosufficiente, ha portato a risultati assai significativi nel caso di giochi come gli scacchi, tradizionalmente considerati uno dei banchi di prova più adatti a misurare le capacità di analisi e scelta di un computer in un contesto suscettibile di una descrizione puramente formale»[8]; tuttavia, aggiungeva ancora Dreyfus[9], quel *cyborg* inteso come 'persona' che agisce nel mondo-della-vita, entrato nella stanza, non riusciva a trovare la scacchiera. In altre parole, se vogliamo usare l'argomento applicandolo ai nostri intenti: l'*uomo artificiale* processa le regole di un linguaggio (quello logico, non ambiguo e *non* referenziale degli scacchi) ma non riesce a usarle se non in un dominio o ambiente affatto ristretto – ovvero in un gioco linguistico determinato, e per di più in

trollare un corpo» (ivi, p. 16) – caratteristiche che gli *smart objects* contemporanei sembrano possedere. In prima approssimazione, si può dire che il problema è relativo all'aspetto intenzionale dell'atto di riferimento che l'AI non sembra possedere, quelle «'regole di ingresso del linguaggio' che, dall'esperienza che abbiamo delle mele, ci portano a pronunciare frasi del tipo 'vedo una mela' e delle regole di uscita dal linguaggio che, da una decisione espressa in forma linguistica (per esempio 'vado a comprare delle mele') ci permettono di passare a azioni che non sono semplicemente di ordine verbale» (ivi, p. 17).

[8] G. Roncaglia, *Computer che copiano: test di Turing, web corpora e filtraggio collaborativo*, in T. Orlandi, a cura, *Per il centenario di Alan Turing, Fondatore dell'informatica*, Atti del Convegno, Roma, Accademia dei Lincei, 2014, pp. 189-201; la cit. è a p. 194.

[9] H.L. Dreyfuss, *What Computers Can't Do*, cit., trad. it. pp. 163 ss.

situazioni prevedibili/calcolabili. Certo, quando Dreyfus criticava i progetti del MIT si era ancora all'alba degli sviluppi degli studi sull'AI; ma già la prospettiva wittengnsteiniana sembrava profilarsi come argomento utile per confutare, proficuamente, le tesi di un approccio 'forte' al tema. Quanto sin qui evidenziato, dunque, rafforza l'ipotesi che alcune forme di vita, emerse nell'ambito dell'AI più recente, siano interpretabili e comprensibili proprio a partire delle riflessioni (pre-digitali) delle *Ricerche filosofiche*. In particolare, è utile rileggere alcuni significativi paragrafi di queste ultime per porli in relazione con le tematiche dell'ormai pluridecennale dibattito sull''intelligenza' dei computer, a partire dagli articoli di Turing scritti negli stessi anni delle *Ricerche* di Wittgenstein e del famoso (per alcuni famigerato) test di Turing, poi ulteriormente problematizzato da Searle nel 1980 nella risposta all'AI classica fornita dall'altrettanto celebre argomento della *Chinese room* ('stanza cinese')[10].

2. *La spinosa (e ambigua) domanda di Turing*

La domanda posta da Turing nel noto articolo del 1950[11], con le sue tre riformulazioni, chiarisce già alcuni punti essenziali delle questioni filosofico-linguistiche implicate nell'interazione uomo-AI (e che è possibile affrontare attraverso una rilettura della nozione di *Lebensform* – 'forma di vita').

[10] Cfr. J.R. Searle, *Minds, Brains and Programs*, in «Behavioral and Brain Sciences», vol. 3, issue 3, September 1980, pp. 417-424; trad. it., *Menti, Cervelli, programmi*, in D. Hofstadter, a cura, *L'io della mente*, Milano, Adelphi, 1985, pp. 341-60.

[11] A.M. Turing, *Computing Machinery and Intelligence*, in «Mind», New Series, vol. 59, n. 236, 1950, pp. 433-460.

> It was suggested tentatively that the question, 'Can machines think?' should be replaced by 'Are there imaginable digital computers which would do well in the imitation game?' If we wish we can make this superficially more general and ask 'Are there discrete-state machines which would do well?' But in view of the universality property we see that either of these questions is equivalent to this, 'Let us fix our attention on one particular digital computer C. Is it true that by modifying this computer to have an adequate storage, suitably increasing its speed of action, and providing it with an appropriate programme, C can be made to play satisfactorily the part of A in the imitation game, the part of B being taken by a man?' [12].

La logica sottesa alle riformulazioni della domanda *Can machines think?* alle quali il test proposto da Turing fornisce, empiricamente, una risposta, configura già una tendenza particolarmente evidente che ritroveremo sia nell'approccio wittgensteiniano alla comunicazione linguistica interumana delle *Ricerche* sia, successivamente, nelle proposte avanzate dalla cosiddetta AI forte a base funzionalista (a partire gli anni Sessanta) ma anche, dalla fine dello scorso secolo, in quella

[12] Turing, *Computing Machinery*, cit., p. 442. [Si è sostenuto provvisoriamente che la domanda "Le macchine possono pensare?" dovrebbe essere sostituita da "sono immaginabili computer digitali che riuscirebbero nel gioco dell'imitazione?" Se vogliamo, possiamo rendere la domanda più generale e chiederci: "Ci sono macchine a stati discreti che riuscirebbero bene?" Ma alla luce della proprietà di universalità, si constata che entrambe le domande sono equivalenti a questa: "Concentriamo la nostra attenzione su un particolare computer digitale C. È vero che, modificando questo computer in modo che disponga di una memoria adeguata, aumentandone opportunamente la velocità d'azione e dotandolo di un programma appropriato, C può riuscire a interpretare in modo soddisfacente la parte di A nel gioco dell'imitazione, laddove la parte di B è interpretata da un uomo?] Trad. degli autori.

versione del funzionalismo che rinuncerà alla prospettiva 'internista' sul mentale e svilupperà sino alle conseguenze più estreme l'idea di un'estensione della mente al di fuori del cervello[13].

Dunque, *Le macchine possono pensare?* – domanda che nella possibile applicazione all'interazione linguistica uomo-macchina possiamo assimilare al titolo di questo articolo: *Le macchine possono significare?* o, meglio, *I programmi significano?* – in effetti è una formulazione affascinante, ma legata a un orizzonte fenomenologico in cui *pensiero* e *significato* sono processi intenzionali originari e privati, (o «in prima persona», per usare l'espressione di Searle), che vengono realizzati fisicamente (sono, cioè, provvisti di una qualche materia e sostanza dell'espressione) ma appaiono descrivibili come interni alla mente – vale a dire risultano verificabili e indagabili pubblicamente solo fondandosi sulle osservabili risposte in *output* a determinati *input* (siano tali *output* comportamentali, «azioni non di ordine verbale» per dirla con Putnam, o semplici rappresentazioni-in-immagini dell'attività cerebrale associata al pensiero).

Invece, riformulando la prima domanda come *È possibile immaginare dei computer digitali capaci di riuscire nel gioco dell'imitazione?* (e ancora, generalizzando, *Esistono macchine a stati discreti che siano in grado di riuscire?*) l'approccio diviene decisamente funzionalista ed esternalista: non si pone più l'accento sulle caratteristiche *interne* del comunicare, bensì sul modo in cui un processo di calcolo sintattico 'imita' un 'motore semantico' sulla base del ruolo causale che esercita[14]. E si noti

[13] M. Di Francesco, G. Piredda, *La mente estesa. Dove finisce la mente e comincia il resto del mondo?*, Milano, Mondadori Università, 2012.

[14] Così Dennett, cit. in Di Francesco, Piredda, *La mente estesa*, cit. p. 52.

che Turing parlava di «gioco dell'imitazione» in un senso non troppo dissimile da quello con cui Wittgenstein si riferiva ai giochi linguistici quali forme di vita (RF §§ 7, 23 e 65-88).

La struttura finale assunta dalla domanda, dunque, non fa che ricondurre ancora una volta un argomento generale a un possibile caso concreto (la base empirica del test, insomma): chiedersi se esista «un particolare computer digitale C» che, se provvisto di una quantità di dati, una velocità di processazione e un programma adeguati possa sostenere con successo la parte di A nel gioco dell'imitazione in cui B è impersonato da un uomo reale, in effetti, significa immaginare che la sussistenza di un'interazione viene ormai valutata esclusivamente su basi fattuali – o di adeguatezza causale –, dando tuttavia per scontato che A sia una macchina autosufficiente non estensibile. Curiosamente, questo è quel che è accaduto per le prime *chatbot*, ma non avviene più nei dispositivi nati negli ultimi decenni dopo lo sviluppo della Rete: essi, infatti, attingono a corpora di dati in linea di principio illimitati per generare e 'confezionare' le risposte date all'utente.

3. *Wittgenstein e l'osservazione senza pensiero*

Ora, passando da Turing a Wittgenstein, due aspetti colpiscono del modo in cui il filosofo viennese affronta il problema del linguaggio:

– il suo rifiuto del *significato* (mentale, privato ma intenzionale[15]) come nozione capace di dar conto degli *usi* del lin-

[15] Sul linguaggio privato e il significato mentale in Wittgenstein cfr. A.J.P. Kenny, *Wittgenstein*, London, Penguin, 1973; trad. it. *Wittgenstein*, Milano, Boringhieri, 1984, pp. 206-34.

guaggio (pubblici, intersoggettivamente costituiti dalle/nelle prassi enunciative), pervicacemente perseguito;

– il conseguente ripudio della prospettiva consistente nel considerare la comunicazione un modo per «trasmettere pensieri» intesi come «processi interni» (RF §§ 304-5), mediante una sorta di 'percorso' che coinvolge attori *umani* ma ricorda il problematico (e decontestualizzato) *circuit de la parole* saussuriano (criticato ad esempio da Roy Harris[16]).

In questo senso, prima ancora di introdurre quello che sembra l'argomento più critico nei confronti di tale rifiuto – ossia la sintassi senza semantica della 'stanza cinese' di Searle – possiamo dire che proprio quest'ultima è anticipata dall'approccio insito nell'analisi wittgensteiniana: il fruttivendolo che interpreta il biglietto «cinque mele rosse» al § 1 delle RF («– Ma che cos'è il significato della parola 'cinque'? – Qui non si faceva parola di un tale significato; ma solo del modo in cui si usa la parola 'cinque'») non differisce in nulla dal Searle dell'esperimento mentale il quale, chiuso nella sua stanza con alcune scatole piene di caratteri cinesi, utilizza le regole di un manuale per rispondere con scelte appropriate di caratteri-simboli a domande giuntegli sottoforma di caratteri-simboli – implementando un programma in assenza di comprensione.

Deriva da ciò la particolare attrattiva dell'approccio delle RF che è bensì osservativo – RF § 66: «non pensare, ma osserva!» – eppure non del tutto meccanicistico e behaviorista, posto che il comportamentismo si confuta da sé quando è messo alla prova del fatto che, nel negare l'occorrere di un'esperienza interna in aggiunta al comportamento osservabile, si nega anche la possibilità di spiegare le relazioni cau-

[16] R. Harris, *The Language Myth*, London, Duckworth, 1981.

sali fra interno ed esterno (così, ad esempio, la credenza non può esser ridotta al comportamento se non postulando l'esistenza di uno stato epistemico di *desiderio* che va spiegato[17]). Wittgenstein insomma riflette criticamente sul behaviorismo metodologico e le sue conseguenze quando afferma, in RF §§ 307-8, che la finzione al di fuori del comportamento è *grammaticale* e, dunque, regolata in base prove ripetute, ma senza poter affermare nulla circa gli stati interni e al tempo stesso senza poterli negare: «dunque dobbiamo negare un processo non ancora compreso [...] E così sembra che abbiamo negato i processi spirituali. E, naturalmente, non li vogliamo negare!» (§ 308).

Il negare la possibilità di dar conto degli stati interni può essere constatato nel caso di alcuni giochi linguistici come, ad esempio, il *leggere*: «non includo nel 'leggere' il comprendere il senso di ciò che viene letto [ma] il convertire in suoni ciò che è scritto» (§ 156). In tale prospettiva, scrive Wittgenstein, un uomo potrebbe addirittura «funzionare come una semplice macchina per leggere» (*ibid.*). La macchina per leggere, però, sembra descrivere esattamente ciò che accade quando attanti dotati di AI interagiscono con soggetti umani.

Alcune delle caratteristiche ritenute essenziali al pensiero (*in primis* linguistico) nell'ambito della filosofia della mente e delle scienze cognitive paiono messe in certo senso 'fuori gioco' dal rifiuto wittgensteiniano del significato-comprensione fondato fenomenologicamente:

– anzitutto, non ha più alcuna importanza che il pensiero esprima un contenuto – posto che il pensiero-*Gedanke* non è più considerato autonomo e ontologicamente distinto

[17] J.R. Searle, *Mind. A Brief Introduction*, Oxford, Oxford University Press, 20024; trad. it. *La mente*, Milano, Raffaello Cortina Editore, 2005, p. 48.

dagli enti materiali e dai processi psicologici (alla Frege, per intenderci). Il modello 'telementale' saussuriano, il cui problema essenziale è quello di «spiegare in che modo i pensieri siano trasferiti da una mente alle altre mediante processi che non siano telepatici»[18], in effetti, è espressione di una tesi del senso comune 'presa di mira' nelle RF, là dove Wittgenstein afferma che è necessario rompere radicalmente «con l'idea che il linguaggio funzioni sempre in un unico modo, serva sempre allo stesso scopo: trasmettere pensieri» (§ 304). Il modello saussuriano riduce infatti il pensiero solo a uno stato interno, di cui tuttavia il funzionalismo può dar conto esclusivamente in termini causali-computazionali;

– in secondo luogo, non ha più senso l'appello al pensare e al parlare quale parte integrante dell'esperienza del soggetto perché – se questo riferimento alla cognizione come comprensione caratterizza l'approccio di Searle, ad esempio – in Wittgenstein e nel funzionalismo computazionale i criteri del riferimento alla coscienza, all'agire in base a ragioni ecc., non sono più considerati *empiricamente* validi ai fini della spiegazione del mentale;

– analogamente, in terzo luogo, il poter parlare di qualcosa di diverso *dal sé*, ovvero l'esser provvisto di intenzionalità oltre che di *agentività* (vale a dire della proprietà di fungere da 'motore' di un'azione, per molti effettivamente constatabile nel funzionamento di *devices* digitali) ci era sembrato un criterio efficace per distinguere l''intelligenza' degli *smart objects* da quella di attori umani[19], ma anche quest'ulti-

[18] R. Harris, *Communication and Language*, in N. Love, ed., *The Foundations of Linguistic Theory. Selected Writings of Roy Harris*, New York, Routledge, pp. 136-50, cit. a pp. 142-3.

[19] R. Finocchi, A. Perri, P. Peverini, *Smart objects in daily life: Tackling the rise of new life forms in a semiotic perspective*, «Semiotica», 235, vol. 1-4, pp. 1-26.

ma caratteristica non è pertinente se il gioco linguistico viene compreso attraverso un approccio osservativo che esclude illazioni su stati interni di qualunque natura. Stati interni che Wittgenstein non intendeva negare, anche se appaiono indiscutibilmente non indagabili a partire da un approccio computazionale il quale, appunto, nega proprio il carattere più 'debole' del pensiero linguistico, ossia la sua natura di esperienza di un soggetto, coerentemente articolata a comporre enunciati e ragionamenti validi perché *significativi* e *compresi* da qualcuno (non, dunque, solo formalmente validi).

4. *AI, devices in Rete, il gioco dell'imitazione e la mente estesa (di nuovo)*

A quanto fin qui esposto vogliamo aggiungere un'ulteriore osservazione. Il 'gioco dell'imitazione' inaugurato dai *devices* connessi alla Rete in grado di processare un linguaggio naturale (*Natural Language Processing*) scardina l'idea della *black box* computazionale autonoma (ossia del computer come cervello). Come è stato fatto notare[20], con i *virtual assistants* il computer ha imparato a 'barare': quando deve reagire a una frase dell'utente, o rispondere a una domanda, il *device* può farlo non solo utilizzando le 'conoscenze' e la base dati di cui già dispone ma anche – e ormai ciò accade inevitabilmente – attingendo a conoscenze e modelli di comportamento linguistico (ancora una volta, però, computati per via di semplice manipolazione sintattica) ricercati opportunamente in rete perché possano essere copiati e 'incollati' nell'interazione.

[20] Roncaglia, *Computer che copiano*, cit.

Questa circostanza rende allora davvero possibile (ancora una volta!) immaginare i *devices* come *menti (non umane) estese*? Le 'credenze' dei dispositivi (in realtà mere informazioni computabili e 'copiabili') sono accessibili in effetti soltanto *disposizionalmente* (ovvero, in termini semiotici, dotate di stati di esistenza virtuali o potenziali prima di esser reperite in Rete) ma allorché le informazioni vengono recuperare e trasformate in occorrenze (attualizzate, realizzate) ciò avviene attraverso un processo computazionale che 'funziona' come la citata macchina per leggere delle RF (§ 156). Ad esempio *Siri*, l'assistente virtuale sviluppato dalla Apple per i suoi prodotti, sinora 'funziona' proprio così (quantunque in un futuro che è già quasi un presente non si limiterà all'*information retrieval* ma provvederà a un'elaborazione generativa dei dati reperiti); dunque cosa distingue, in prospettiva, un *vocal assistant* dal taccuino di Otto nel noto esperimento mentale di Clark e Chalmers[21]? La risposta sembrerebbe essere '*nulla*' – almeno se capiamo che, effettivamente, questi dispositivi si 'muovono' nella Rete senza comprendere né significare ma solo computando su simboli: la supposta 'mente estesa' di *Siri* non è (e non può essere) quella di un interlocutore in dialogo con noi, e neppure si 'fonde' meramente con la nostra mente biologica di utenti dotati di cervello e sensorialità – in modo del tutto simile alle estensioni 'oggettuali' quale il taccuino (una forma di intenzionalità derivata del soggetto, anche se provvista di agentività intrinseca). Al contrario, pensare a *Siri* o *Alexa* (assistente domestico sviluppato da Amazon) come a delle *menti* rappresenta una *nostra* proiezione di intenzioni e attività cognitive sul *device*, proiezione a sua volta

[21] A. Clark, D. Chalmers, *The Extended Mind*, in «Analysis», 58, 1, pp. 7-19.

fondata su una presunzione fenomenologica indotta proprio dalla *Lebensform*: l'idea di poter *dialogare* con *Siri*.

Siri insomma (o *Alexa*, ma anche e in misura ancor più radicale *ChatGPT*) *non ha credenze*, non *significa*: copia e risponde ai nostri *input* con *output* sintatticamente coerenti non interpretati dal programma C (che imita A, un soggetto umano assente) bensì dall'altro utente umano B.

Pensiamo a quel che accade nei casi in cui interagiamo con un *vocal assistant* utilizzando predicati epistemici che introducono, come noto, la controversa nozione fregeana di *ungerade Bedeutung*, problematica già nella formulazione che Frege ne aveva dato[22]: il significato di un enunciato introdotto da predicati epistemici non è un valore di verità estensionale cui il suo *Sinn* rinvia, bensì *il pensiero stesso* (e noi diremo, abbracciando un'epistemologia non ontologica post-fregeana, il *contenuto mentale intenzionale*) che l'enunciato in isolamento esprime. Ora, Siri è ad oggi programmato in modo da *bypassare* il predicato epistemico della domanda che gli viene rivolta in ogni circostanza in cui il pensiero retto dal predicato è un enunciato che contenga simboli reperibili (e 'copiabili' in quanto tali) dal *corpus* di dati della Rete (ovvero enunciati *fattuali*, che non rinviano a stati interni non descrivibili in termini di proposizioni verificabili):

> *Siri (o Alexa) che ore sono? / Sono le 10:32 di mattina* è così assimilabile *a Siri (o Alexa) credi di sapere che ore sono? / Sono le 10:32.*

Ma se proviamo a sostituire l'interrogazione con un wittengsteiniano riferimento all'ineludibile 'privatezza' di un

[22] P. Casalegno, *Il paradigma di Frege*, in Santambrogio, M., a cura, *Introduzione alla filosofia analitica del linguaggio*, Roma-Bari, Laterza, pp. 3-40, in part. pp. 32-7.

enunciato come *Siri provi dolore?* o *Siri credi di provare dolore?* – riformulando l'asserzione come enunciato complesso con un predicato epistemico – la risposta del *device* sarà in entrambi i casi *Quale?* (o *Questo è quello che ho trovato*), accompagnata da una serie di *link* dalla Rete in forma di lista che rimandano a una serie di centri di terapia del dolore localizzati sulla base del posizionamento dell'utente[23]. Ancora, alla domanda relativa a stati interni 'privati' dell'utente umano – quali *Siri credi che io provi dolore?* – Siri, sino a pochi mesi fa, rispondeva (a sua volta 'copiando' un predicato epistemico come *credere* che, di fatto, 'non comprende') *Non credo di capire. Come altro posso aiutarti?* mentre oggi 'copia' la risposta alla domanda diretta al dispositivo, equiparando *credi di provare dolore* (a sua volta 'ridotto' a *provi dolore*) e *credi che io provi dolore* (e rispondendo, dunque, anche a quest'ultima domanda con *Quale?* seguito da relative indicazioni dei medesimi centri di terapia del dolore). In altre parole, il dispositivo non è in grado di 'gestire' l'opposizione tra la prima e la terza persona, naturalmente non in quanto meri *token* linguistici processabili ma in relazione alla differenza irriducibile fra i contenuti *non fattuali* esprimibili mediante l'occorrenza di tali deittici. Analogamente, quando uno di noi ha rivolto a *Siri* per la prima volta la domanda *credi che io sia simpatico?* (è avvenuto 'solo' quattro anni fa!) la risposta del *vocal assistant* di allora – sublime reviviscenza di un'antinomia sepolta nella storia della filosofia del linguaggio – era stata: *Non credo*

[23] Quelli citati sono gli esiti di una verifica empirica sull'*iPhone* e sul dispositivo *Alexa* compiuta dagli scriventi nel mese di luglio 2023. Come si è osservato, è verosimile che gli *upload* occorsi nella programmazione abbiano contribuito alla possibilità che i *devices* forniscano già oggi, o in un prossimo futuro, risposte diverse (e nel caso di *Siri* le osservazioni si riferiscono anche a una replica dell'interazione svolta nel momento in cui scriviamo [luglio 2024]).

di avere credenze; oggi *Siri* risponde come faceva in passato alla domanda sul dolore privato dell'utente: *Non credo di capire. Come altro posso aiutarti?* (conservando il predicato epistemico). Sembra davvero che il *machine learning*, in questo caso, abbia prodotto i suoi frutti, posto che l'elaborazione del *device* replica (e simula) proprio un 'suo' processo epistemico-interpretativo evidentemente insussistente (o comunque inverificabile da un punto di vista fenomenologico). Alla domanda *Siri credi di essere simpatica?*, d'altro canto, l'assistente vocale reagisce ancora una volta come se si trattasse di un'asserzione *priva* del predicato epistemico, e priva altresì della 'forza' di direttivo/interrogativo – atto che, pragmaticamente, esige una risposta –, considerando l'enunciato una mera asserzione fattuale (quale *Siri sei simpatica*) e 'rispondendo' coerentemente da un punto di vista pragmatico, *Mi fa piacere!* o *Lo prendo come un complimento.*

5. *Conclusioni: una forma di vita riflessa?*

Come ha giustamente notato Searle, in un ripensamento critico del più volte citato argomento della *Chinese room*[24], l'assenza di semantica non esaurisce la problematicità del modo di funzionare del dispositivo se posto in relazione alla mente biologica umana: *anche la computazione e la sintassi, in effetti, sono relative all'osservatore*, «perché non ci sono in natura computazioni intrinseche o originarie»[25] ma solo derivate (ovvero attribuite dagli agenti umani). Torna, ancora una volta, l'emblematico esempio della 'macchina per leggere

[24] Searle, *Mind. A Brief introduction*, cit., trad. it. p. 83.
[25] *Ibidem.*

umana' di Wittgenstein, la quale legge «a voce alta e correttamente, senza prestare attenzione a ciò che legge» (§ 156): in *Siri* questa figura viene semplicemente riproposta come voce di sintesi che processa – *senza comprendere la sua stessa computazione e sintassi* – una serie di simboli 'copiati' dalla Rete.

Tutte le attività che possiamo prendere in esame, allora, rinviano indiscutibilmente a forme di vita (RF § 23) nelle quali le intelligenze artificiali *per noi* si 'comportano' come agenti soggettivi (quantunque *non* intenzionali).

Eppure, agli utenti umani che partecipano al gioco dell'imitazione tutti questi dispositivi appaiono in grado di realizzare interazioni genuine, e finiscono persino per esser dotati di una qualche intenzionalità (derivata), instaurando con gli utenti rapporti giudicati produttivi (e *sensati*, ma solo e sempre nell'ottica degli agenti umani). Questo effetto-riflesso appare estremamente contagioso e manipolatorio, in effetti, e lo dimostrano le stesse parole recentemente pronunciate da Chomsky, Robert e Watumull contro *The False Promise of ChatGPT*[26]: condannando quella forma di *machine learning* per la sua «moral indifference born of unintelligence», la si accusa addirittura di mostrare «something like the banality of evil: plagiarism and apathy and obviation». Si contestano le prestazioni del *device*, certo, ma riferendosi ad esso come si trattasse di un agente 'inumano' affetto dalla terribile anestesia morale descritta dalla Arendt!

Ecco, allora, che l'approccio radicalmente anti-internista del Wittgenstein delle *Ricerche* ci fa comprendere come tutto ciò sia possibile – e, forse, necessario – al funzionamento delle nostre forme di vita con l'AI.

[26] N. Chomsky, I. Robert, J. Watumull, *Noam Chomsky: The False Promise of ChatGPT*, in «The New York Times», Monday 8, 2023, in rete al *link*: *https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html.*