

L'intelligenza artificiale come assistente e tutor linguistico: alcune ipotesi di lavoro

Stefania Spina
Università per Stranieri di Perugia

Abstract

Il contributo fornisce un quadro sintetico dei potenziali campi di applicazione delle tecnologie emergenti legate all'intelligenza artificiale nell'apprendimento linguistico. Dopo avere definito tali tecnologie come una potentissima "capacità di agire", si sofferma ad analizzare alcune ricerche che hanno già valutato gli effetti dell'intelligenza artificiale sulla didattica delle lingue, sia dal lato dei docenti, in cui può rivelarsi un utilissimo assistente linguistico, sia dal lato degli apprendenti, in cui può svolgere rilevanti funzioni di tutor linguistico. Il contributo si propone di sottolineare opportunità, sfide e rischi dell'intelligenza artificiale e di promuoverne un uso consapevole.

1. L'intelligenza artificiale e la didattica delle lingue: minaccia o opportunità?

Dal novembre 2022, quando la società OpenAI ha lanciato ChatGPT, un sistema di intelligenza artificiale (IA) in grado di imitare un'interazione linguistica umana, il suo uso in ogni ambito della società e della conoscenza si è diffuso rapidamente, e insieme si è sviluppato un intenso dibattito sulle implicazioni, positive o negative, di questo uso generalizzato (per il contesto italiano, cfr. Cicero 2023). L'enorme impatto di questo tipo di IA è legato a tre sue caratteristiche fondamentali:

- è una IA *generativa*: utilizza un modello di *machine learning* per apprendere da enormi quantità di dati e contenuti creati dall'uomo (si parla infatti di *Large Language Models*, o LLM). Da questi dati appresi genera nuovi contenuti, che, a seconda della complessità del modello sul quale si basano, sono in grado di generare testo, immagini, audio o video;
- è una IA *conversazionale*: i LLM sono sistemi basati su interfacce in grado di dialogare con l'utente;
- è una IA *dotata di un'interfaccia amichevole* e semplice da usare per l'utente, e resa disponibile a chiunque per lo sviluppo di applicazioni.

Il dibattito ha investito anche il campo dell'apprendimento linguistico. Chi considera l'IA come una minaccia, anche in questo campo, si pone domande del tipo: se un telefonino sarà in grado di parlare e tradurre tutte le lingue in tempo reale, che senso avrà imparare una lingua straniera? Sarà un'attività riservata a pochi specialisti? E che ruolo rimarrà al docente di lingue? Le domande sembrano legittime, anche considerando che il ruolo del docente di lingue è già da tempo "minacciato" da app gratuite per telefonini che forniscono a chi vuole apprendere una lingua straniera input graduato da usare in autonomia, che il software è in grado di correggere all'infinito. A maggior ragione, l'IA, con le sue straordinarie potenzialità di rielaborare le risposte dell'apprendente, potrà rendere superfluo l'intervento del docente?

Sul versante opposto, chi considera l'IA come un'opportunità per innovare i metodi didattici, ne intravede la possibile integrazione con la didattica tradizionale, in contesti educativi sia formali che informali, per sviluppare forme di didattica nuove. Ad esempio, dal lato docente, l'opportunità può consistere nel delegare all'IA specifiche attività che prevedano l'elaborazione e il controllo di grandi quantità di dati, come ad esempio la produzione di materiale didattico o l'individuazione degli errori degli apprendenti. Dal lato studente, invece, l'IA offre l'opportunità di sviluppare interazioni dialogiche con "partner conversazionali", da cui ricevere in tempo reale feedback personalizzato.

2. IA: la capacità di agire separata dall'intelligenza

Un primo passaggio fondamentale per orientarsi all'interno del mondo al tempo stesso promettente e minaccioso dell'IA è quello di definirla: che cos'è realmente l'IA? Un aspetto importante per la sua definizione è che l'elaborazione di enormi quantità di dati, la velocità di calcolo e il perfezionamento nel tempo degli algoritmi sono tutte operazioni statistiche, dotate di efficacia sulla struttura formale dei testi ma non sul loro significato. L'IA, dunque, agisce sugli aspetti formali e non su quelli semantici dei dati che elabora. È, quindi, una forma di *solving agency* (Floridi), una capacità di azione mai sperimentata prima, perché è in grado di imparare e di migliorare il proprio comportamento senza dover essere *intelligente*. Con l'IA, dunque, abbiamo a disposizione un sistema potentissimo, dotato di una straordinaria capacità di azione; quindi, di eseguire compiti lunghi e impegnativi, senza essere di fatto dotato di intelligenza. L'IA separa la capacità di agire dall'intelligenza: *fa* senza capire.

Data questa fondamentale premessa, la domanda che come docenti e ricercatori dobbiamo porci, è: quella tra IA e didattica delle lingue è un'interazione possibile?

3. Un'interazione possibile?

La domanda se la sono posta, ad esempio, i docenti e ricercatori del British Council, che nel 2023 hanno pubblicato uno studio dal titolo *Artificial intelligence and English language teaching: Preparing for the future* (Edmett et al.), in cui hanno analizzato le ricerche già pubblicate sul tema e hanno sottoposto un questionario a 1.348 docenti di inglese L2 di 118 paesi diversi. L'analisi delle ricerche pubblicate ha evidenziato che i benefici maggiori derivati dall'interazione tra IA e didattica dell'inglese L2 si registrano in particolare nel campo della produzione orale (attraverso l'uso di sistemi basati sull'IA per la pratica della pronuncia, che registrano e forniscono feedback alla pronuncia degli apprendenti) e della produzione scritta (ad esempio attraverso sistemi che aiutano gli apprendenti a produrre testi grammaticalmente e sintatticamente corretti). Il questionario rivolto ai docenti ha invece evidenziato che gli strumenti basati sull'IA usati nella didattica dell'inglese servono in ordine decrescente a creare materiali didattici (57%), a favorire la pratica dell'inglese da parte degli apprendenti (53%), a creare e pianificare lezioni (43%), a correggere l'inglese degli apprendenti (33%) e a valutare gli apprendenti (23%).

Il lavoro del British Council, anche se risale a un paio d'anni fa, mette in evidenza come non solo un'interazione tra IA e didattica linguistica sia possibile, ma sia già largamente in atto, specialmente in alcune aree. Ciò che diventa di primaria importanza, in un'epoca in cui sperimentare nuovi strumenti tecnologici nella didattica è una pratica corrente, è disporre di ricerche sistematiche e rigorose per valutare l'efficacia e l'impatto dell'IA nella pratica didattica concreta. Lo scopo dei prossimi paragrafi è quello di portare come esempio alcune delle ricerche già esistenti, nella duplice prospettiva dei docenti e degli apprendenti.

4. L'IA per il docente di lingue

Alcuni studi recenti hanno analizzato l'efficacia dell'IA per i docenti, in alcuni compiti che riguardano la valutazione delle competenze, l'individuazione degli errori degli apprendenti e la preparazione di materiali da usare in classe.

Per quanto riguarda la valutazione, l'area di ricerca applicata denominata *Automated essay scoring* (AES) è attiva da un cinquantennio, quindi da molto prima dello sviluppo dell'IA. Il suo obiettivo è da sempre servirsi della tecnologia per valutare testi scritti di apprendenti sulla base di criteri linguistici predefiniti (lessicali, morfosintattici, ecc.). Tra i vantaggi che una valutazione assistita dalla tecnologia può portare, oltre al risparmio di tempo, c'è la riduzione dell'incidenza della inevitabile componente soggettiva nella valutazione. Con lo sviluppo dell'IA, si è cercato di sfruttare i LLM nella pratica di AES. Mizumoto e Eguchi è uno studio di ampio respiro mirato appunto a valutare questa integrazione. Lo studio prende in esame 12.100 testi inglesi scritti da

persone con 11 diverse L1 che hanno sostenuto l'esame TOEFL per la certificazione dell'inglese L2 nel 2006 e nel 2007. Gli stessi testi erano già stati valutati da valutatori umani: ogni testo era assegnato a uno dei tre livelli iniziale, medio, o avanzato. Questo vastissimo campione di testi è stato dato da valutare a ChatGPT-3, con un prompt contenente una lista di criteri linguistici di valutazione riguardanti il lessico, la grammatica e la coesione testuale. La valutazione dell'IA è stata poi confrontata con quella dei valutatori umani, per misurarne l'accuratezza. I risultati mostrano che ChatGPT raggiunge un discreto livello di accuratezza, ma non un accordo perfetto con le valutazioni umane. Di conseguenza, può diventare un'importante risorsa di supporto ai valutatori umani, ma non in grado di sostituirli del tutto. Lo studio evidenzia anche l'importanza del *fine-tuning*: addestrando in modo specifico il sistema di IA, si possono ottenere risultati migliori. Un'altra area di applicazione dell'IA che ha delle potenzialità importanti è quella dell'identificazione automatica degli errori, o *Grammatical error detection* (GED). L'individuazione degli errori è un compito che riveste una grande importanza in diversi ambiti, come la didattica delle lingue, il language testing e in generale la ricerca acquisizionale. Fioravanti et al. ha messo a confronto quattro sistemi di IA (ChatGPT-4o, Copilot, Gemini and Llama3) con due gruppi di annotatori umani: parlanti nativi di italiano e apprendenti di italiano L2. Sia i sistemi di IA che gli annotatori umani hanno annotato 11 testi di livello B1 tratti dal corpus CELI, che contiene testi prodotti per la certificazione di italiano per i livelli B1-C2 da parlanti di diverse L1 (Spina et al. 2022). Il task, proposto in un prompt molto dettagliato, era quello di individuare nei testi otto tipi di errori morfosintattici riguardanti articoli, preposizioni, verbi e nomi; gli errori potevano essere di selezione, aggiunta o omissione. I risultati dello studio hanno mostrato che i sistemi IA sono molto meno accurati nell'individuazione degli errori anche dei parlanti L2, e raggiungono un livello di accuratezza non ancora sufficiente perché un docente possa delegare loro la correzione degli errori nei testi scritti da apprendenti. Inoltre, la prestazione dei LLM dipende anche dal tipo di errore: ad esempio, è stata riscontrata una maggiore accuratezza negli errori di omissione di articoli e preposizioni. Anche in questo caso, addestrando in modo mirato il sistema attraverso operazioni di *fine-tuning* si possono ottenere prestazioni migliori.

Una delle aree di applicazione più produttive dell'IA relativamente alla didattica linguistica è quella della produzione di materiali didattici. In questo caso, lo scoglio più difficile da superare è legato al livello di competenza. Il quesito che si pongono alcuni studi recenti è dunque: i sistemi di IA sono in grado di produrre materiali didattici adatti a specifici livelli di competenza? Uno studio svolto in collaborazione tra l'università di Stanford e Duolingo (Malik et al.) parte dal presupposto che il problema da superare è la sostanziale ambiguità delle scale *Quadro Comune Europeo di Riferimento* (QCER): mentre infatti i livelli non adiacenti (es. B1 e C1) sono sufficientemente differenziati gli uni dagli altri, i livelli adiacenti, specialmente in testi brevi, hanno sovrapposizioni e sono difficilmente differenziabili. Problemi di questo tipo si riscontrano anche con sistemi non basati sull'IA, come MALT-IT2, uno strumento in grado di misurare in modo automatico la difficoltà di un testo italiano e di assegnarlo ad un livello nella fascia B1-C2 (<https://lol.unistrapg.it/malt/>). Malik et al. dimostra che, senza un addestramento specifico, i LLM raggiungono un livello di accuratezza non ancora sufficiente perché un docente possa delegare loro la produzione di testi conformi ai diversi livelli QCER. In particolare, lo studio rivela una forte sovrapposizione tra i livelli del range B1-C1. Viceversa, addestrando in modo specifico il sistema con un *fine-tuning* mirato ai livelli QCER, dando vita a un ulteriore sistema chiamato CaLM (CEFR-Aligned Language Model) si ottengono risultati molto promettenti.

5. L'IA come tutor linguistico per l'apprendente di lingue

Dal lato degli apprendenti, l'IA ha moltissime potenzialità, in particolare nelle aree della facilitazione dell'interazione nella lingua target, nella disponibilità di feedback personalizzato e di strumenti di riferimento supportati dall'IA, come i dizionari.

L'IA è uno strumento prezioso per chi apprende una L2 in primo luogo perché è in grado di promuovere l'interazione autentica nella lingua target, attraverso partner conversazionali. Da

diverso tempo esistono applicazioni che consentono la pratica linguistica orale individuale col supporto di interlocutori virtuali generati dall'IA, con cui è possibile dialogare su temi e a livelli di competenza scelti dagli utenti. Sui testi prodotti gli apprendenti ricevono feedback, sia a livello grammaticale che fonetico. Se queste applicazioni sono molto promettenti e stanno avendo un rapido sviluppo, la ricerca non ha ancora prodotto risultati affidabili sulla loro efficacia: Xiao et al., ad esempio, ne riconosce le potenzialità ma conclude che su questi *conversational agents* "research in the L2 learning context is still emergent and scarce".

Un altro contributo rilevante dell'IA, in particolare per le attività autonome degli apprendenti, è legato al ruolo che alcune piattaforme di IA possono svolgere come supporto e guida all'apprendimento. *NotebookLM* di Google, ad esempio, consente di caricare testi e di dialogare con un chatbox generato dall'IA ottenendo feedback personalizzato su quei testi, mirato ad esempio alla correzione degli errori, alla riformulazione delle frasi, al miglioramento delle scelte lessicali o sintattiche, ecc. Oltre ai vantaggi di poter contare su feedback linguistico adatto alle proprie esigenze, strumenti di questo tipo promuovono una modalità di apprendimento come interazione e scoperta guidata i cui benefici sono noti, anche se studi mirati alla valutazione dell'efficacia di queste piattaforme di IA come supporto e guida all'apprendimento non sono ancora disponibili. Hanno già cominciato ad apparire, invece, ricerche che valutino l'efficacia dell'IA come dizionari, quindi come strumenti di riferimento per il lessico usato nella lingua target. Questi studi cercano di fornire risposte alla seguente domanda: l'IA è in grado di integrare (o sostituire) il dizionario, una risorsa di riferimento usata dagli apprendenti all'interno della classe o in modo individuale? Ptasznik et al. confronta l'efficacia del *Longman Dictionary of Contemporary English* (LDOCE) usato da apprendenti polacchi di inglese L2 con ChatGPT4, sia in compiti di produzione che di comprensione. Per limitarci alla sola comprensione, nel test previsto dallo studio ai partecipanti era richiesto di comprendere 20 frasi in inglese, concentrandosi in particolare su una parola poco comune, e poi di tradurla in polacco, servendosi del dizionario o dell'IA. Le traduzioni sono quindi state valutate, giungendo alla conclusione che ChatGPT4 è in grado di fornire supporto lessicale efficace agli apprendenti, per due ragioni principali: perché è semplice da consultare, e fornisce un supporto lessicografico contestualizzato; inoltre, è uno strumento che coinvolge attivamente gli apprendenti in interazioni dialogiche. L'IA, dunque, può essere in grado di aiutare gli apprendenti a risolvere problemi di tipo lessicale sia a livello produttivo che ricettivo. Queste conclusioni possono avere anche interessanti sviluppi in campo lessicografico: ad esempio, per il costituendo *Dizionario delle collocazioni italiane per apprendenti* (DICI-A; Spina et al. in preparazione), dopo un'ampia fase di sperimentazione le definizioni delle 11.000 collocazioni inserite nel dizionario sono state create col supporto dell'IA.

6. Conclusioni: opportunità, sfide e rischi

Nel 2023, una tavola rotonda sull'IA nel campo dell'educazione organizzata dall'UNESCO concludeva che "these emerging technologies present immediate – as well as far reaching – opportunities, challenges and risks to education systems" (UNESCO). Le opportunità sono legate alla collaborazione attiva con i sistemi di IA per sostenere il percorso di insegnamento/apprendimento linguistico. Dal punto di vista dei docenti, il vantaggio principale risiede nella possibilità di delegare all'IA compiti lunghi e gravosi, basati sull'elaborazione di testi e di dati, ad esempio per la creazione di materiali didattici, test, attività e per il supporto all'individuazione degli errori e alla valutazione. Per gli apprendenti, l'IA consente di ricevere feedback immediato e personalizzato, sviluppare interazioni dialogiche con partner conversazionali, ricevere supporto linguistico nelle attività di apprendimento. Per docenti e apprendenti, si aprono infinite possibilità di usare efficacemente l'IA in modo creativo. Le sfide sono legate alla necessità di promuovere, nei docenti di lingua come negli apprendenti, un'interazione consapevole con l'IA, nell'ottica di un loro uso critico e responsabile. Una delle sfide principali è imparare a dialogare con i sistemi di IA: apprendere cioè ad usare i prompt più efficaci e a valutare l'accuratezza delle risposte. Il *prompt engineering* è infatti alla base dell'interazione con i sistemi di IA, e richiede

l'adozione di un linguaggio preciso e appropriato, basato su una valutazione continua delle variabili da considerare, partendo dal presupposto che solo a un input efficace corrisponde un output efficace. Questo implica ovviamente la necessità di una formazione e di un aggiornamento continui. Il rischio principale è invece quello di recepire in modo acritico e assumere a priori come vero, corretto, appropriato e utile tutto ciò che un sistema di IA ci offre in risposta ai nostri quesiti, anche didattici: le frequenti allucinazioni dei sistemi di IA dimostrano invece che tutto va sempre verificato e vagliato in modo consapevole e critico. In assenza di ciò, si rischia concretamente di azzerare la complessità, e instillare (o rafforzare?) negli studenti l'idea che con l'IA sia tutto semplice e alla portata di un clic, anche apprendere una lingua: basta chiedere a GPT. Concludendo, l'IA può essere utilizzata come assistente o tutor linguistico con vantaggi ancora tutti da esplorare, entro un perimetro delimitato dall'assunto fondamentale che l'IA è capacità di azione più che intelligenza: una risorsa preziosa a cui attingere in modo al tempo stesso consapevole e creativo. Nel ruolo di assistente linguistico, all'IA possono essere demandati compiti specifici, con ocularità, in alcuni casi avendola preventivamente addestrata; come tutor linguistico, l'IA consente di valorizzare l'apprendimento linguistico come scoperta progressiva attraverso il dialogo e l'interazione. In entrambi i casi, è indispensabile un'attività costante di formazione e una connessione molto stretta con la ricerca.

Riferimenti bibliografici

Cicero, Francesco. "L'italiano delle intelligenze artificiali generative". *Italiano LinguaDue*, 15 (2023): 733-761.

Floridi, Luciano. "AI as Agency Without Intelligence: On ChatGPT, Large Language Models, and Other Generative Models". *Philosophy & Technology*, 36 (2023).

Edmett, Adam, Ichaporia, Neenaz, Crompton, Helen e Crichton, Ross. *Artificial intelligence and English language teaching: Preparing for the future*. British Council. British Council, 2024.

Fioravanti, Irene, Forti, Luciana e Spina, Stefania. "Automatic Error Detection: Comparing AI vs. Human Performance on L2 Italian Texts". In *Proceedings of the 10th Italian Conference on Computational Linguistics (CLiC-it 2024)*, 366–372, Pisa, Italy, 2024.

Malik, Ali, Mayhew, Stephen, Piech, Chris e & Bicknell, Klinton. "From Tarzan to Tolkien: Controlling the Language Proficiency Level of LLMs for Content Generation". *arXiv* (2024).

Mizumoto, Atsushi e Eguchi, Masaki. "Exploring the potential of using an AI language model for automated essay scoring". *Research Methods in Applied Linguistics*, 2 (2023).

Ptasznik, Bartosz, Wolfer, Sasha e Lew, Robert. "A Learners' Dictionary Versus ChatGPT in Receptive and Productive Lexical Tasks". *International Journal of Lexicography*, 37 (2024): 322–336.

Spina, Stefania, Fioravanti, Irene, Forti, Luciana, Santucci, Valentino, Scerra, Angela e Zanda, Fabio. "Il corpus CELI: una nuova risorsa per studiare l'acquisizione dell'italiano". *Italiano Lingua Due*, 14 (2022): 116–138.

Spina, Stefania et al. (in preparazione). *DICI-A. Dizionario delle collocazioni italiane per apprendenti*.

UNESCO. *Global Education Monitoring Report 2023: Technology in Education-A Tool on Whose Terms?*, 2023.