

L'INTELLIGENZA ARTIFICIALE GENERATIVA A SUPPORTO DELLE ABILITÀ DI SCRITTURA IN L2/LS NEL CONTESTO UNIVERSITARIO: UNA SCOPING REVIEW

Mariapaola Paita, Sofia Romagnoli¹

1. INTRODUZIONE

Negli ultimi anni, l'evoluzione delle tecnologie basate sull'intelligenza artificiale (IA) ha inciso profondamente sul panorama educativo, favorendo l'integrazione di sistemi intelligenti a supporto dei processi di insegnamento e apprendimento (Mattioda *et al.*, 2023). Le applicazioni dell'IA in ambito educativo (*Artificial Intelligence in Education*, AIED) sono ad oggi impiegate per offrire un'esperienza didattica personalizzata mediante la creazione automatica di materiali e ambienti digitali interattivi e adattivi che permettono il monitoraggio costante dei progressi di apprendimento e la restituzione di feedback automatizzati e calibrati sui bisogni e sulle caratteristiche degli studenti (Chiu *et al.*, 2023; Wang *et al.*, 2024). Sistemi come gli *Intelligent Tutoring Systems* per l'apprendimento adattivo e l'analisi predittiva degli obiettivi di apprendimento tramite la profilazione degli studenti (Long, Siemens, 2011; Mousavinasab *et al.*, 2021), o le tecnologie specializzate in *Automated Written Feedback* (AWF), tra cui gli *Automated Essay Scoring* (AES) e gli *Automated Writing Evaluation* (AWE) stanno contribuendo a trasformare le pratiche educative, offrendo nuove opportunità di personalizzazione, monitoraggio e valutazione a tutti i livelli dell'istruzione (Crompton, Burke, 2022, 2023).

Grazie ai progressi nel campo del trattamento automatico del linguaggio (*Natural Language Processing*, NLP), del *machine learning* e, più recentemente, degli algoritmi di *deep learning* (Schmidt, Strasser, 2022), le potenzialità dell'IA nel promuovere un apprendimento «more personalised, flexible, inclusive, and engaging» (Holmes, Luckin, 2016: 11) hanno suscitato una crescente attenzione anche in ambito edulinguistico. In particolare, con l'avvento dell'IA Generativa (IAG) l'interesse già marcato per gli strumenti di IA nell'educazione linguistica ha subito una rapida accelerazione. Tra le sue diverse applicazioni, la capacità dei *Large Language Models* (LLM) di generare contenuti multimodali e simulare conversazioni tramite input testuali ha ampliato in modo significativo le prospettive per l'insegnamento e l'apprendimento delle lingue (Kasneji, 2023). Numerosi studi evidenziano come essi possano assumere il ruolo di tutor virtuali, in grado di promuovere l'autonomia, la motivazione e la riflessione metacognitiva degli apprendenti (Chiu *et al.*, 2023; Lai *et al.*, 2023), o di agenti conversazionali a supporto dello sviluppo e del potenziamento delle competenze linguistiche e strategiche, tra cui la mediazione, l'interazione, e la produzione scritta (tra i tanti contributi in merito si vedano Su *et al.*, 2023; Young, Shishido, 2023; Katelhön, 2025).

¹ Università di Parma.

La presente *scoping review* è il risultato del lavoro congiunto dei due autori. Mariapaola Paita è responsabile dei §§ 1.2, 1.3 e 4 e Sofia Romagnoli ha redatto i §§ 1.1 e 3. Sono da attribuirsi a entrambi gli autori i §§ 1, 2 e 5.

Sebbene l'implementazione degli strumenti basati sulla IAG per la didattica delle lingue sia ancora in una fase esplorativa, il potenziale pedagogico associato all'utilizzo di queste tecnologie nei contesti di apprendimento linguistico è evidenziato da Katelhön (2025: 478), che sottolinea come esse rappresentino

un potenziale enorme per trasformare, e soprattutto, per personalizzare l'esperienza di apprendimento delle lingue straniere. Gli strumenti basati sull'IA, come chatbot, assistenti virtuali e piattaforme di apprendimento adattivo, permettono di simulare interazioni reali e offrire feedback immediato agli/alle apprendenti. Inoltre, grazie all'elaborazione del linguaggio naturale (NLP), questi strumenti possono analizzare e correggere errori, suggerire miglioramenti e adattare i contenuti alle esigenze individuali. Questo approccio potenzia l'acquisizione linguistica attraverso un'esposizione costante e interattiva alla lingua target.

1.1. *La scrittura in L2/LS: approcci teorici e implicazioni didattiche*

La scrittura in lingua seconda o straniera (L2/LS) si configura come un'attività produttiva che prevede l'integrazione di conoscenze linguistiche, strategie cognitive e metacognitive, nonché la gestione di variabili affettive. La sua natura intrinsecamente riflessiva la rende una competenza complessa che implica la trasformazione di concetti astratti in testi coerenti e coesi attraverso un processo di progettazione e riordino delle idee, realizzazione e revisione del testo (Hedge, 2011). La distinzione tra lingua seconda (L2) e lingua straniera (LS), introdotta in ambito italiano da Balboni (2006), si fonda sulla diversa natura dell'esposizione alla lingua: la L2 è appresa in un contesto in cui la lingua è presente nell'ambiente sociale e educativo dell'apprendente, mentre la LS è appresa in un contesto privo di input autentici, spontanei e continui.

Nel contesto dell'apprendimento linguistico, la scrittura assume un valore epistemico: non si limita a riflettere il livello di competenza acquisito ma costituisce uno strumento per l'elaborazione e la ristrutturazione delle conoscenze linguistiche. L'apprendente è chiamato a integrare conoscenze lessicali e grammaticali, garantire coesione e coerenza testuale, organizzare logicamente i contenuti e adottare un registro comunicativo adeguato al contesto. Inoltre, deve essere in grado di formulare ipotesi linguistiche, monitorare la propria produzione e intervenire autonomamente attraverso processi di autocorrezione e revisione. La produzione scritta in L2/LS si colloca quindi in una prospettiva di apprendimento orientata al processo e che valorizza il ruolo dell'errore, dell'autoregolazione e del feedback; in cui la lingua è attivamente costruita e adattata agli scopi comunicativi e cognitivi richiesti dal compito. Essa rappresenta uno spazio in cui l'apprendente costruisce attivamente la propria interlingua, intesa come un sistema linguistico autonomo e provvisorio, distinto sia dalla lingua madre dell'apprendente sia dalla lingua target, che emerge durante il processo di acquisizione e si evolve gradualmente attraverso l'esposizione e l'uso della lingua (Selinker, 1992).

È attraverso la scrittura, infatti, che l'apprendente attiva processi di *noticing* (Schmidt, 1990) e *hypothesis testing* (Swain, 1993) che gli permettono di identificare le proprie lacune, formulare ipotesi linguistiche e monitorare la propria produzione e revisione. Queste fasi trovano riscontro anche nella *Skill Acquisition Theory* sviluppata da DeKeyser (2020), per cui l'attività di produzione scritta favorisce la trasformazione delle conoscenze dichiarative in conoscenze procedurali, colmando la distanza tra le competenze passive e attive. All'interno della cornice teorica proposta dalla *Sociocultural Theory* (Vygotsky, 1978) la scrittura si configura come un'attività durante la quale l'apprendimento risulta mediato da

forme di *scaffolding* interattivo, fornito da un interlocutore esperto o un agente artificiale, che permettono all'apprendente di agire all'interno della Zona di Sviluppo Prossimale (ZPD).

In ambito accademico, l'esercizio della scrittura risulta complicato dalla necessità di possedere non soltanto competenze linguistiche avanzate, ma anche la capacità di conformarsi alle convenzioni retoriche e a stili definiti. Si tratta di un'attività, soprattutto all'interno di un contesto L2/LS, più complessa e sono numerosi gli studi che ne sottolineano tanto l'importanza, quanto le difficoltà (Jubier *et al.*, 2024; Mezzadri, 2024). La motivazione di un apprendente assume, in questo contesto, un ruolo determinante nel sostenere il processo di scrittura. La costruzione di un *Ideal L2 Self*, ovvero l'immagine futura di sé come utente competente della lingua (Dörnyei, 2009), può rafforzare la motivazione strumentale e intrinseca. Analogamente il concetto di *self-efficacy* elaborato da Bandura (1997), secondo cui la percezione di autoefficacia incide significativamente sul comportamento dell'apprendente, risulta fondamentale, così come i filtri affettivi, che possono ostacolare l'elaborazione degli input necessari all'acquisizione (Jubier *et al.*, 2024; Mezzadri, 2024).

Il feedback diventa dunque fondamentale nel configurare la scrittura non come esecuzione di un compito ma come parte integrante di un processo formativo, in cui l'apprendente è guidato a rielaborare e migliorare progressivamente la propria produzione. Hattie e Timperley (2007: 1, enfasi nell'originale) lo definiscono come una «information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding. [...] Feedback thus is a consequence of performance». Winstone e Carless (2019) propongono un approccio interattivo e co-costruito al feedback, in cui lo scambio tra emittente e ricevente diventa un'occasione per promuovere la consapevolezza e l'autonomia dell'apprendente. In quest'ottica, il feedback diventa il mezzo attraverso il quale è possibile colmare il divario tra la prestazione dell'apprendente e gli obiettivi formativi attesi. Il suo scopo è, dunque, quello di agire come facilitatore del processo di autocorrezione, al fine di elaborare strategie di revisione e riscrittura e favorendo la riflessione sul testo. Il processo di revisione da parte dell'apprendente, inteso come «a sequence of changes in a composition - changes which are initiated by a check and occur continually throughout the writing of the work» (Sommers, 1980: 380) assume quindi un ruolo fondamentale nel processo di scrittura.

In un contesto di insegnamento linguistico tradizionale della LS/L2 è l'insegnante a decidere quale tipologia di feedback applicare e secondo quali modalità di correzione. Patthey-Chavez *et al.* (2004), riprendendo la distinzione proposta da Olson e Raffeld (1987), distinguono tra *surface-level feedback* e *content-level feedback*. I primi si concentrano su aspetti formali, fornendo indicazioni sulla correttezza grammaticale o sulle scelte lessicali, mentre i secondi vengono usati per evidenziare problematiche legate ai contenuti, alle strategie retoriche e alla coesione del testo. Nonostante i benefici riconosciuti, nelle pratiche didattiche tradizionali non sempre è possibile offrire un riscontro formativo individualizzato a causa dell'elevato numero di studenti per classe o la mancanza di risorse temporali e formative. A ciò si somma la tendenza a ricorrere prevalentemente a una specifica tipologia di feedback rispetto ad un'altra, che può portare ad enfatizzare troppo la correttezza del testo, a scapito del contenuto, e viceversa (Lee, 2008).

Alla luce di queste considerazioni risulta giustificato il crescente interesse verso la IAG come tecnologia a supporto della didattica. In particolare, i LLM ne rappresentano una componente particolarmente rilevante per l'ambito edulinguistico, in quanto consentono di affrontare tali criticità migliorando la qualità e l'efficacia del feedback fornito agli apprendenti.

1.2. I sistemi di Automated Writing Evaluation per la produzione scritta

Un ambito particolarmente rilevante delle applicazioni basate sull'IA nell'educazione linguistica è costituito dai sistemi di valutazione automatica della produzione scritta (AWE) (cfr. Ranalli, Hegelheimer, 2022). Questi strumenti rappresentano una tipologia di IA che si avvale di algoritmi di NLP e analisi semantica latente (*Latent Semantic Analysis*, LSA) per analizzare in maniera automatizzata le produzioni scritte degli apprendenti e fornire un feedback immediato e personalizzato (Shi, Aryadoust, 2024). A differenza dei sistemi di AES, progettati principalmente per la valutazione sommativa, gli AWE sono impiegati soprattutto in contesti formativi per favorire il potenziamento delle abilità di scrittura attraverso un feedback mirato e personalizzato della produzione scritta dell'apprendente (Stevenson, Phakiti, 2019).

L'origine degli AWE risale agli anni '60, con lo sviluppo del *Project Essay Grader* basato su misurazioni come la lunghezza testuale e la presenza di punteggiatura, ma privo di capacità semantiche (Page, 2003). Con l'introduzione della LSA negli anni '90 e il progresso tecnologico nel campo del NLP, sono emersi strumenti di AWE più complessi e sofisticati, in grado di offrire un feedback mirato su una vasta gamma di aspetti formali e strutturali del testo, tra cui la grammatica, il lessico, la morfosintassi e l'organizzazione testuale (Deane, 2013). Tra gli AWE attualmente più diffusi si annoverano gli E-rater (per esempio, Criterion, Grammarly e Pigai) e gli *Intelligent Tutoring Systems*, come *The Writing Pal*, specializzati per guidare gli apprendenti nelle diverse fasi del processo di scrittura (Roscoe *et al.*, 2014).

Nonostante il potenziale educativo degli AWE nel consolidare le abilità di scrittura, la letteratura scientifica ne evidenzia alcuni limiti (cfr. Stevenson, Phakiti, 2019). Innanzitutto, il feedback generato si concentra maggiormente sugli aspetti formali del testo, quali l'accuratezza grammaticale, l'ortografia e la punteggiatura, mentre dimensioni più complesse della scrittura, come lo sviluppo delle idee, la coerenza testuale e lo stile autoriale, risultano ancora poco esplorate (Burstein *et al.*, 2013; Deane, 2013). Inoltre, la natura prevalentemente predefinita e priva di interazione dinamica del feedback può limitare lo sviluppo di competenze metacognitive da parte degli apprendenti (Hyland, Hyland, 2006).

Un'ulteriore criticità riguarda l'impostazione normativa di molti AWE, rafforzando una visione prescrittiva della lingua che può penalizzare il repertorio linguistico, il contesto educativo e socioculturale e gli obiettivi comunicativi dell'apprendente (Vojak *et al.*, 2011; Stevenson, Phakiti, 2019). Ciò risulta particolarmente rilevante se si considera che la maggior parte degli AWE sono stati originariamente progettati per parlanti nativi e pertanto, nonostante il loro ampio utilizzo nei contesti di English as a Second/Foreign Language (Wang *et al.*, 2013), non riflettono pienamente le esigenze e caratteristiche specifiche degli apprendenti di una L2 o LS.

1.3. L'Intelligenza Artificiale Generativa a supporto delle abilità di scrittura

Negli ultimi anni si è registrato un crescente interesse per le tecnologie basate sulla IAG, e in particolare per le potenzialità dei LLM. Grazie alla loro natura conversazionale e generativa, questi strumenti consentono un'interazione più adattiva e personalizzata con l'utente, rendendoli potenziali risorse per l'apprendimento linguistico (per una panoramica sulle funzionalità e sulla natura dei LLM si rimanda a Yang *et al.*, 2024). Sebbene la ricerca sia ancora in fase iniziale, i primi contributi evidenziano come ChatGPT possa avere un impatto positivo sulla motivazione, l'autonomia e il miglioramento delle competenze linguistiche degli apprendenti di una L2 o LS (Albadarin

et al., 2024; Teng, 2024). Parallelamente, gli studi segnalano anche criticità, tra cui il rischio di una dipendenza eccessiva, che rappresenta un potenziale ostacolo allo sviluppo del pensiero critico e della creatività, e problematiche legate alla privacy e all'integrità accademica (Crompton *et al.*, 2024; Pérez *et al.*, 2025).

Un ambito di ricerca attualmente in espansione riguarda l'implementazione dei LLM, in particolare ChatGPT, a supporto della produzione scritta, con una prevalenza di studi incentrati sull'apprendimento dell'inglese come LS in ambito accademico (Ling Jen, Salam, 2024). In particolare, la letteratura riporta un utilizzo ricorrente di ChatGPT durante il processo di scrittura nel ruolo di assistente durante la pianificazione e di valutatore automatico in fase di revisione, offrendo un feedback personalizzato sugli aspetti formali del testo (Li, 2024). Tra i principali vantaggi sono menzionati, per esempio, l'ampliamento del repertorio lessicale e un aumento dell'accuratezza grammaticale (Al-khreshehm, 2024). Allo stesso tempo, i timori legati al rischio di plagio e alla violazione dell'integrità accademica hanno portato la ricerca ad interrogarsi sull'efficacia degli *AI detectors* per distinguere i testi generati dalla IAG da quelli scritti dall'uomo, al fine di garantire un uso etico e responsabile dei LLM (Chaka, 2024).

La letteratura scientifica sull'uso dell'IAG in ambito edulinguistico risulta pertanto in costante espansione e numerosi studi in merito iniziano a esplorarne il potenziale nel promuovere le abilità di scrittura di apprendenti di una L2 o LS nei contesti accademici. Tuttavia, le revisioni sistematiche disponibili tendono a concentrarsi principalmente sull'efficacia degli strumenti di AWE (cfr. Ding *et al.*, 2024; Huang *et al.*, 2024) o sulle tecnologie basate sull'IA prima dell'avvento della IAG (cfr. Shi, Aryadoust, 2024; Pérez *et al.*, 2025). Inoltre, le ricerche condotte sui LLM per promuovere l'apprendimento linguistico esplorano i benefici e le sfide legate all'uso di ChatGPT durante la produzione scritta degli apprendenti nei contesti di EFL/ESL (cfr. Al-khreshehm, 2024; Ling Jen, Salam, 2024; Teng, 2024).

In questo scenario, la presente *scoping review* si propone di offrire una visione aggiornata e sistematica del panorama della ricerca attuale sulla IAG per promuovere le abilità di scrittura in L2/LS nei contesti di istruzione superiore. Particolare attenzione è dedicata all'utilizzo dei chatbot basati sui LLM, al fine di identificare le tipologie e le strategie di implementazione adottate, il ruolo educativo loro attribuito e le dimensioni dell'apprendimento linguistico maggiormente investigate durante la produzione scritta mediata da queste tecnologie.

2. METODOLOGIA

Alla luce della natura esplorativa degli obiettivi di ricerca di questo contributo, è stata condotta una *scoping review* per far emergere le evidenze disponibili nella letteratura scientifica riguardante l'uso delle tecnologie basate sulla IAG a supporto delle abilità di scrittura in L2/LS². A tal proposito, si è adottato il *framework* metodologico proposto da Arksey e O'Malley (2005) e ampliato da Levac *et al.* (2010), che prevede un processo sistematico di revisione della letteratura articolato in cinque fasi principali:

1. definizione delle domande di ricerca;
2. identificazione degli studi rilevanti;
3. selezione degli studi;

² Benché l'intelligenza artificiale generativa (IAG) comprenda diverse tipologie di strumenti per la generazione di contenuti, nel contributo il termine è impiegato in senso operativo per riferirsi alle applicazioni basate su modelli linguistici di grandi dimensioni (LLM), più strettamente connesse alle attività di scrittura.

4. estrazione dei dati;
5. sintesi e presentazione dei risultati.

In questa sezione saranno descritte le quattro fasi del processo di revisione, mentre per la sintesi e presentazione dei risultati si rimanda alla Sezione 4.

2.1. Definizione delle domande di ricerca

Gli obiettivi che guidano la presente *scoping review* sono stati tradotti nelle seguenti domande di ricerca, al fine di esplorare sistematicamente il panorama della letteratura sulla IAG per la produzione scritta in L2/FL:

- D1. Qual è il panorama della ricerca sugli strumenti basati sulla IAG per la produzione scritta in L2/FL in ambito accademico in termini di anno di pubblicazione, paese, partecipanti e approccio metodologico?
- D2. Quali sono i principali quadri teorici o *framework* concettuali adottati negli studi?
- D3. Quali tipologie, con quali modalità di interazione e in quale contesto educativo vengono impiegati gli strumenti basati sulla IAG per promuovere le abilità di scrittura in L2/FL?
- D4. Qual è il ruolo educativo attribuito agli strumenti basati sulla IAG nel supportare le abilità di scrittura in L2/FL?
- D5. Quali competenze e dimensioni legate alla produzione scritta vengono esplorate?

2.2. Identificazione degli studi

Gli studi rilevanti sono stati identificati attraverso una strategia articolata in due fasi principali. Diverse combinazioni di parole chiave e operatori booleani sono stati innanzitutto testati per affinare la stringa di ricerca in relazione agli obiettivi della revisione, garantendo una copertura quanto più esaustiva e mirata possibile della letteratura scientifica disponibile (Arksey, O'Malley, 2005; Levac *et al.*, 2010). Questa fase preliminare ha permesso la definizione di una stringa di ricerca coerente con gli obiettivi della revisione e articolata nella seguente *query*:

(feedback OR “evaluat*” OR “assess*”) AND (“writ*”) AND (“Large Language Model*” OR “LLM” OR “ChatGPT” OR “GPT-3.5” OR “GPT-4” OR “GPT-4o” OR “Bard” OR “Gemini” OR “Claude” OR “Mistral” OR “AI chatbot*” OR “AI-generated text” OR “neural language model*” OR “AI language model*” OR “language model” OR “ai” OR “artificial intelligence” OR “language generation model*” OR “conversational AI”) AND (“language*” OR “linguistic” AND (“education*” OR “foreign” OR “learn*” OR “second” OR “l2” OR “FL” OR “EFL” OR “acquisition”).

L'identificazione degli studi è stata condotta il 15 marzo 2025 nelle banche dati Scopus, ERIC ed EBSCOHost. La ricerca è stata ristretta ai titoli, agli abstract e alle parole chiave e sono stati esclusi gli ambiti disciplinari distanti dal campo linguistico-educativo, tra cui il settore medico, giuridico, economico, agrario e delle scienze ambientali. Queste scelte metodologiche hanno consentito di mantenere un equilibrio tra ampiezza della revisione e rilevanza dei contributi selezionati, garantendo un'ampia copertura della letteratura scientifica potenzialmente pertinente e riducendo il rischio di includere studi non coerenti con i campi dell'educazione linguistica e dell'AIED. Infine, sono stati inclusi soltanto articoli in lingua inglese, pubblicati dal 2022 al 2025 e sottoposti a un processo di revisione

in doppio cieco, al fine di aumentare la qualità degli studi inclusi e la loro pertinenza rispetto agli obiettivi di ricerca (Levac *et al.*, 2010; Daudt *et al.*, 2013).

Questa fase ha permesso l'identificazione di 3.003 studi, che sono stati inclusi nelle fasi successive di *screening* e selezione.

2.3. Selezione degli studi

La selezione degli studi rilevanti è avvenuta in due fasi. In un primo momento, il numero complessivo degli studi identificati nelle banche dati ($n = 3.003$) è stato importato nel software di gestione bibliografica Zotero per la rimozione dei duplicati ($n = 241$). I contributi rimanenti ($n = 2.762$) sono stati quindi sottoposti a un processo di *screening* dei titoli e degli abstract condotta sulla base dei criteri di inclusione ed esclusione precedentemente definiti dai ricercatori e riportati nel dettaglio nella Tabella 1. Durante questa prima fase è stato adottato un approccio inclusivo: gli studi il cui titolo o abstract non specificavano chiaramente la tipologia di IA utilizzata o i risultati ottenuti sono stati inclusi nelle fasi successive di selezione, al fine di evitare l'esclusione prematura di ricerche potenzialmente rilevanti. Ciò ha portato all'identificazione di 102 studi, di cui 8 sono stati successivamente esclusi a causa dell'impossibilità di accedere al testo completo.

Tabella 1. *Criteri di inclusione ed esclusione durante la fase di screening degli studi selezionati*

Criteri di inclusione	Criteri di esclusione
Articoli su rivista e sottoposti a un processo di revisione in doppio cieco.	Monografie, capitoli di monografie, editoriali e commentari, atti di convegno.
Articoli pubblicati a partire dal 2022.	Articoli pubblicati prima del 2022.
Articoli scritti in lingua inglese.	Articoli scritti in lingua diversa dall'inglese.
Il testo completo degli articoli deve essere accessibile.	Il testo completo dell'articolo non è accessibile.
Ricerche empiriche che presentano: risultati relativi ad aspetti della produzione scritta degli apprendenti; una descrizione chiara ed esaustiva della metodologia, degli obiettivi e delle domande di ricerca.	Revisione sistematiche e narrative della letteratura, analisi bibliometriche, ricerche non empiriche.
Studi che prevedono l'uso di LLM (o applicazioni di IAG basate su LLM) durante attività di produzione scritta degli apprendenti.	Studi che utilizzano applicazioni di IA diverse dai LLM.
L'uso del LLM deve avere come obiettivo principale il miglioramento delle abilità di scrittura degli apprendenti.	L'uso del LLM non ha come obiettivo principale il miglioramento delle abilità di scrittura (ad es., studi sulla modalità d'uso, su più abilità e/o competenze non linguistiche).
Studi condotti nell'ambito dell'apprendimento di una L2/LS.	Studi sull'apprendimento della lingua madre (L1).
Studi condotti in contesti universitari.	Studi condotti in contesti scolastici o prescolastici.

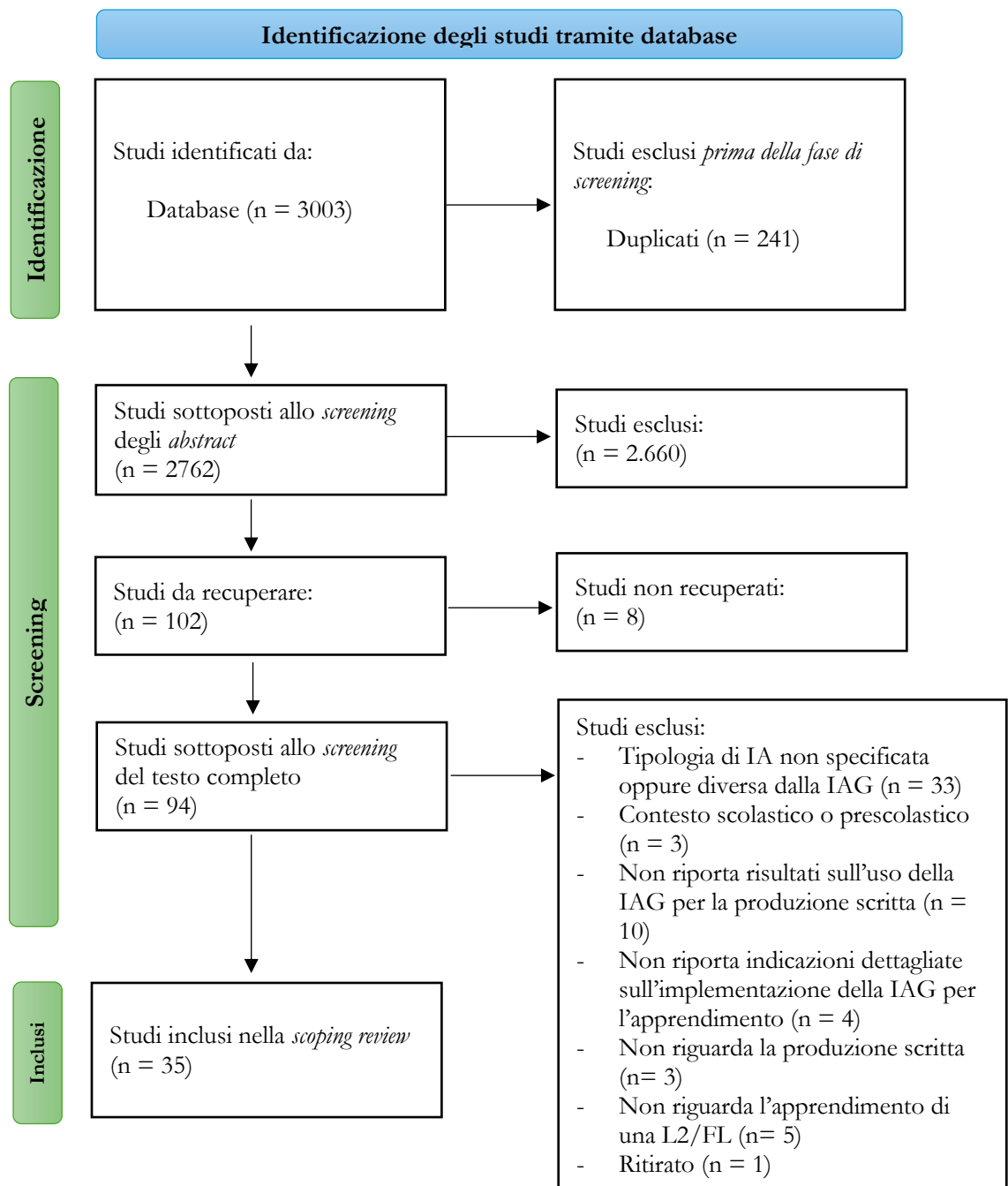
L'intero processo di revisione è stato caratterizzato da un approccio iterativo per garantire un livello costante di trasparenza e coerenza nelle fasi di *screening* e selezione degli studi (Arksey, O'Malley, 2005). A tal proposito, i ricercatori si sono confrontati regolarmente nelle fasi iniziali, intermedie e finali della selezione per risolvere i dubbi emersi dallo *screening* e perfezionare i criteri di inclusione ed esclusione, dove necessario (Levac *et al.*, 2010). In particolare, tutti gli articoli inclusi nella prima fase ($n = 94$) sono stati sottoposti a un secondo *screening* da parte dei ricercatori. Il testo completo di ciascuno studio è stato analizzato e valutato in base ai criteri di inclusione ed esclusione culminato nell'esclusione degli studi che non specificavano la tipologia di IA utilizzata, non riguardavano le applicazioni di IAG basate sui LLM per l'apprendimento di una L2/LS, non avevano come obiettivo principale il miglioramento delle abilità di scrittura in contesti accademici, o non soddisfacevano i requisiti metodologici. Il risultato finale del processo di selezione è terminato con l'inclusione di 35 contributi rilevanti. Una panoramica complessiva del processo di identificazione e selezione degli studi rilevanti è riportata nel dettaglio nella Figura 1.

2.4. Mappatura dei dati estratti dagli studi

Il corpus finale dei 35 contributi che soddisfano pienamente i criteri di eleggibilità sono stati oggetto di un'analisi volta a mappare in maniera sistematica le informazioni pertinenti agli obiettivi della *scoping review*. Per facilitare l'organizzazione, la codifica e la sintesi dei dati, i ricercatori hanno elaborato una griglia di estrazione in formato Excel, successivamente sottoposta a una fase preliminare di *piloting*, al fine di testarne l'efficacia e di verificarne l'adeguatezza rispetto alle domande di ricerca prestabilite. Le incertezze emerse durante l'estrazione sono state discusse tra i ricercatori fino al raggiungimento di un consenso unanime. In conseguenza di ciò, la griglia è stata aggiornata iterativamente per garantire una mappatura dei dati completa, coerente e allineata con le finalità della *scoping review* (Levac *et al.*, 2010). La griglia di estrazione ha pertanto permesso la raccolta sistematica e coerente delle informazioni relative:

1. ai dati bibliografici (anno di pubblicazione, paese di riferimento);
2. al disegno di ricerca (quadro teorico, approccio metodologico, caratteristiche e numero dei partecipanti);
3. alla tipologia di intervento didattico e alla lingua target;
4. al ruolo educativo svolto dalla IAG per promuovere le abilità di scrittura, con particolare attenzione alle applicazioni utilizzate, al feedback sulla produzione scritta e alla modalità di interazione con lo strumento.

Figura 1. Diagramma di flusso PRISMA per l'identificazione e la selezione degli studi rilevanti (Page et al., 2021)



3. RISULTATI

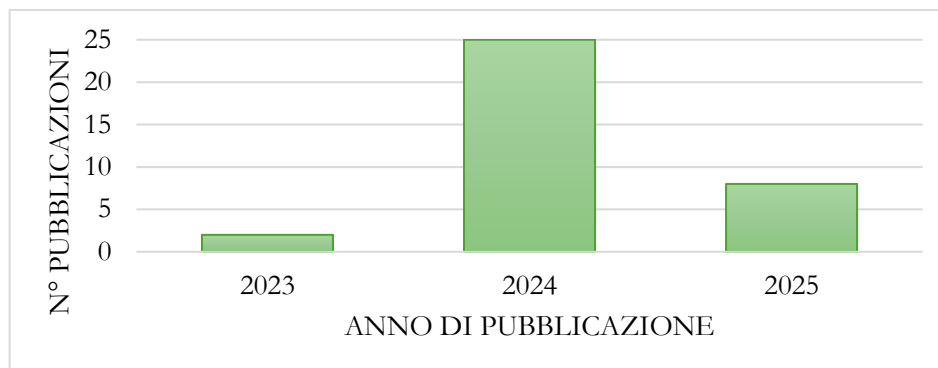
In questa sezione sarà presentata l'analisi e la sintesi dei dati estratti dal processo di mappatura delle informazioni contenute nel corpus finale degli studi, al fine di evidenziare il panorama della ricerca sulla IAG applicata alla produzione scritta in L2/LS e il suo ruolo nel consolidamento delle abilità di scrittura degli apprendenti.

3.1. Panorama della ricerca

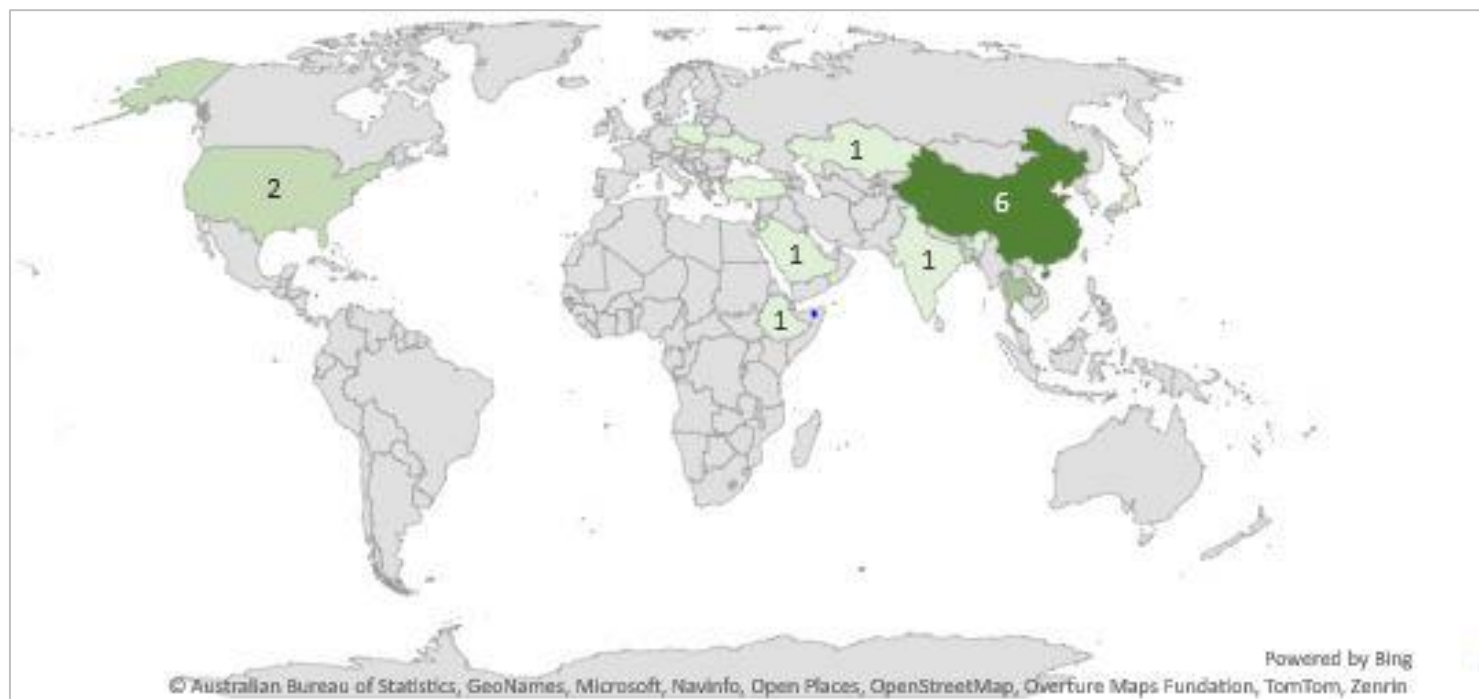
3.1.1. Distribuzione temporale e geografica delle pubblicazioni

La scrittura accademica assistita dalle applicazioni basate sulla IAG rappresenta un campo di ricerca caratterizzato da una marcata concentrazione cronologica e una geografia scientifica sbilanciata. La maggior parte degli studi sono stati pubblicati nel 2024 (n=25), nel primo trimestre del 2025 (n=8) si registra un *trend* analogo o potenzialmente superiore, mentre solo una minoranza è riferibile al 2023 (n=2) (Figura 2). Tale distribuzione conferma il recente aumento di interesse da parte della ricerca, verosimilmente legato alla crescente disponibilità e accessibilità degli LLM a partire dal 2023, in corrispondenza del rilascio pubblico e dell'ampia diffusione di strumenti come ChatGPT.

Figura 2. Distribuzione temporale del numero di pubblicazioni



I dati relativi alla distribuzione geografica riflettono l'interesse globale del mondo della ricerca, sebbene sia rilevabile un netto squilibrio nella distribuzione dei contributi (Figura 3). Una predominanza significativa della letteratura è prodotta nel continente asiatico, in particolare in Cina (n=6), Thailandia (n=3), Bangladesh (n=2), Macau (n=2), Taiwan (n=2), Giappone (n=1), Giordania (n=1), Kazakistan (n=1), Vietnam (n=1), Hong Kong (n=1), Arabia Saudita (n=1), Turchia (n=1), India (n=1) e Corea del Sud (n=1). Dal continente africano proviene un contributo (Etiopia n=1), e alcuni dall'Europa: Paesi Bassi (n=1), Polonia (n=1), Ucraina (n=1), Repubblica Ceca (n=1). Il continente americano è rappresentato dagli USA (n=2).

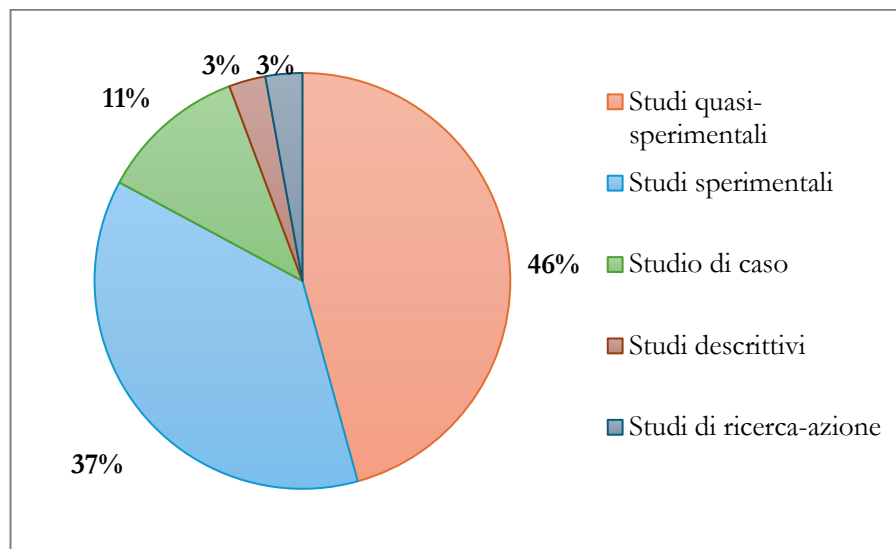
Figura 3. *Distribuzione geografica del numero di pubblicazioni*

3.1.2. *Disegno di ricerca, approcci metodologici e partecipanti*

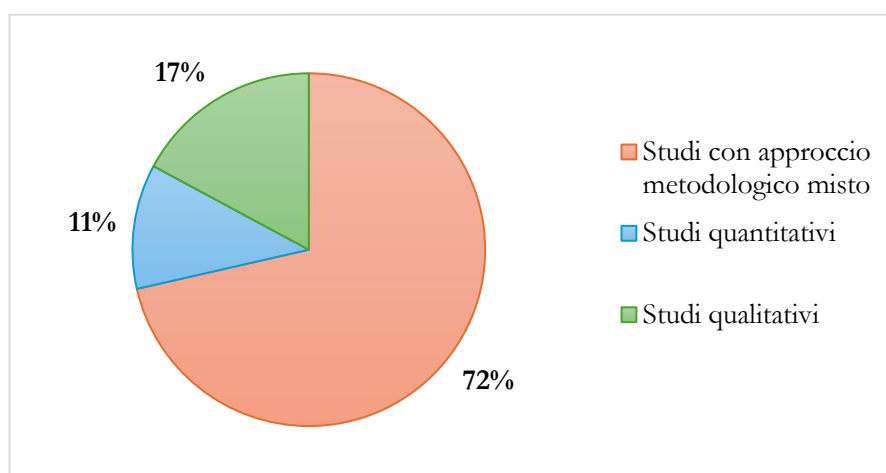
L'analisi degli studi inclusi mostra una marcata inclinazione verso disegni di ricerca di tipo sperimentale, suggerendo un forte interesse verso la valutazione sistematica dell'efficacia didattica della IAG nella scrittura in L2/LS. Come illustrato nella Figura 4, prevalgono gli studi quasi-sperimentali (46%) e sperimentali (37%), mentre risultano meno frequenti gli studi di caso (11%), le ricerche descrittive e di ricerca-azione partecipativa, entrambe corrispondenti al 3% dei contributi. Questo quadro riflette l'esigenza, sempre più diffusa nella comunità scientifica, di fornire evidenze empiriche sull'efficacia degli strumenti basati sul LLM nella didattica della scrittura.

La prevalenza di disegni sperimentali e quasi-sperimentali risponde alla necessità di isolare variabili, misurare l'impatto degli interventi e generalizzare i risultati (Cohen *et al.*, 2017). Questi studi, condotti in ambienti artificialmente controllati e con compiti predefiniti, tendono a privilegiare condizioni di somministrazione standardizzate e temporaneamente limitate, che spesso non riflettono le pratiche autentiche, le dinamiche interattive e le variabili culturali proprie dei contesti reali di apprendimento linguistico. La dimensione ecologica e situata dell'apprendimento, che caratterizza l'uso della lingua in contesti scolastici, accademici o informali, risulta essere pertanto marginalizzata.

La scarsa presenza di approcci partecipativi, etnografici o interpretativi nei contributi esaminati evidenzia inoltre una limitata attenzione alle modalità concrete di interazione con strumenti basati sul LLM, mentre elementi come la negoziazione discorsiva, il posizionamento dell'apprendente, la componente affettiva e la mediazione culturale rimangono ai margini del discorso scientifico. In una prospettiva glottodidattica, tale asimmetria metodologica rischia di ridurre l'esperienza di scrittura mediata da IA a una mera prestazione misurabile, trascurando il suo valore come pratica sociale e culturale.

Figura 4. *Distribuzione degli studi rispetto al disegno di ricerca adottato*

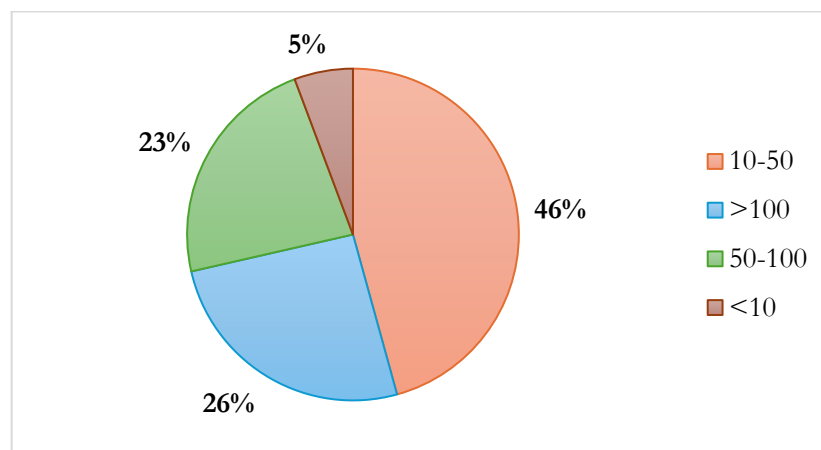
Questa distribuzione eterogenea è riscontrabile anche negli approcci metodologici adottati. Come rappresentato in Figura 5, negli studi analizzati risulta privilegiato un approccio a metodi misti (72%), rispetto ad approcci qualitativi (17%) e quantitativi (11%). Questo orientamento metodologico evidenzia che ad essere preferita è una strategia che permetta una triangolazione dei dati, utile per cogliere la complessità che l'impatto di strumenti emergenti ha nei contesti educativi. In particolare, il ricorso a un approccio di ricerca a metodi misti permette di coniugare l'efficacia misurabile dell'uso di applicazioni basate su LLM, con gli aspetti più soggettivi, quali la percezione e le esperienze degli apprendenti, fondamentali in un contesto L2/LS, dove il successo dell'apprendimento dipende anche da variabili affettive e comportamentali (Biju *et al.*, 2024).

Figura 5. *Distribuzione degli studi in base all'approccio metodologico adottato*

La dimensione del campione costituisce un indicatore altrettanto rilevante per comprendere lo stato attuale del panorama della ricerca sull'utilizzo della IAG nei contesti di apprendimento linguistico. Come riportato nella Figura 6, la maggior parte degli studi analizzati coinvolge campioni compresi tra i 10 e i 50 partecipanti (46%). Seguono gli studi

con un campione di dimensioni medio-grandi, rispettivamente di oltre 100 partecipanti (26%) e compresi tra i 50 e i 100 soggetti (23%). Si rilevano infine due casi con un campione inferiore a 10 partecipanti (5%). La tendenza degli studi ad impiegare campioni di dimensioni medio-piccole, all'interno di singoli corsi o in contesti istituzionali circoscritti, riflette la natura ancora esplorativa e sperimentale degli interventi didattici supportati da LLM. Il maggiore coinvolgimento di campioni ristretti riduce inoltre la possibilità di generalizzare i risultati ottenuti a popolazioni più ampie e rappresentative.

Figura 6. Distribuzione degli studi in base alle dimensioni del campione dei partecipanti



3.1.3. Contesto teorico delle ricerche

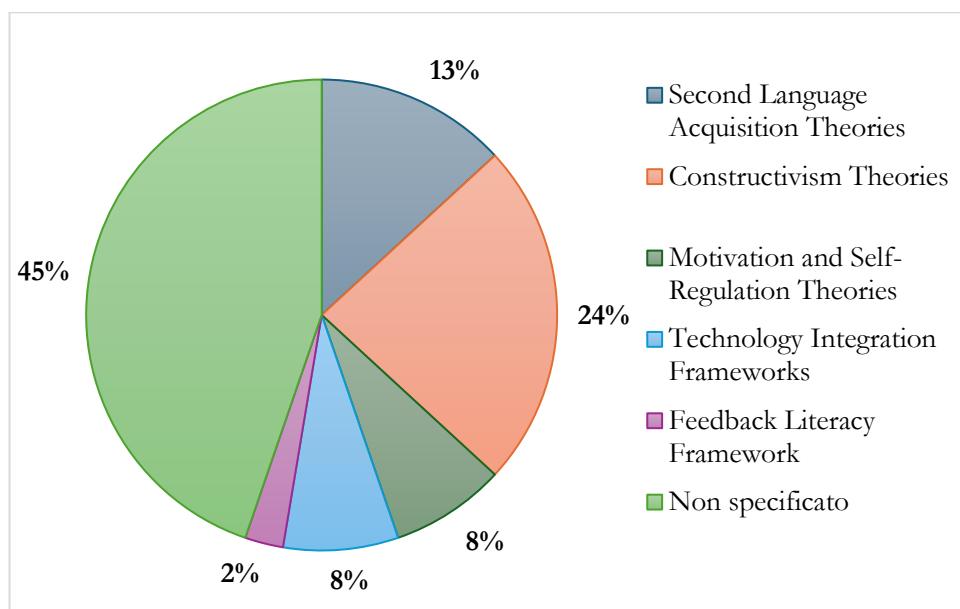
Un ulteriore aspetto che caratterizza il panorama della ricerca relativa all'implementazione delle tecnologie basate sulla IAG per la produzione scritta in L2/LS riguarda i riferimenti teorici adottati negli studi analizzati. Nella Figura 7 è possibile osservare come nel 45% degli studi non venga esplicitato alcun inquadramento teorico a supporto dell'uso della IAG per l'apprendimento linguistico. Questo dato è in linea con le più recenti revisioni sistematiche della letteratura sull'AIED, che sottolineano una scarsa attenzione alle implicazioni didattiche o all'inquadramento teorico-educativo degli strumenti basati sull'IA, a fronte di un orientamento prevalente verso lo sviluppo e la prestazione tecnica degli strumenti proposti (cfr. Chen *et al.*, 2020a, 2020b; Ouyang *et al.*, 2022).

Tra gli studi che adottano esplicitamente *framework* concettuali emerge una preferenza per gli approcci di matrice costruttivista (24%), che valorizzano l'apprendimento linguistico come processo attivo, situato e orientato alla co-costruzione del significato (Powell, Kalina, 2009). Da un lato, gli studi si ispirano al costruttivismo sociale, facendo riferimento alla *Sociocultural Theory* e al concetto di ZPD (Vygotsky, 1978). In questa prospettiva, l'interazione dinamica tra l'apprendente e il *chatbot* durante l'attività di scrittura è concepita come una forma di *scaffolding* interattivo e il feedback generato dai LLM assume una funzione mediatrice, promuovendo il processo di apprendimento all'interno di un ecosistema digitale sociale e collaborativo, anche in assenza di un interlocutore umano³. Dall'altro lato, si rimanda al costruttivismo cognitivo (Piaget, 1953) e all'idea che

³ Nel contributo, utilizzeremo il termine *chatbot* per indicare solamente gli agenti conversazionali basati su *Large Language Models* (LLM), appartenenti all'ambito dell'Intelligenza Artificiale Generativa (IAG) e progettati per simulare conversazioni in linguaggio naturale.

la conoscenza deve essere costruita attivamente dall'apprendente attraverso l'interazione con l'ambiente. In quest'ottica, le capacità conversazionali dei LLM possono stimolare la rielaborazione personale, la riflessione cognitiva e la ristrutturazione interna delle conoscenze durante il processo di scrittura.

Figura 7. Distribuzione degli studi in base al quadro teorico adottato



Parallelamente, il 13% dei contributi adotta teorie provenienti dal campo della Second Language Acquisition al fine di investigare come l'implementazione della IAG durante la scrittura in L2/LS incida sullo sviluppo dell'interlingua dell'apprendente. Tra i riferimenti più ricorrenti figurano la *Noticing Hypothesis* (Schmidt, 1990), che sottolinea l'importanza di notare consapevolmente gli input linguistici come condizione necessaria per attivare il processo di acquisizione linguistica, l'*Interaction Hypothesis* (Long, 1981), che attribuisce un ruolo centrale all'interazione per lo sviluppo delle competenze linguistiche, e la *Skill Acquisition Theory* (DeKeyser, 2020), che riconosce nel graduale processo di automatizzazione attraverso la pratica guidata una condizione necessaria all'acquisizione linguistica. Gli studi che si fondano su queste teorie pongono particolare enfasi sul feedback generato dai *chatbot* in fase di revisione, proponendo un confronto diretto tra i testi prodotti dagli apprendenti e prototipi target, con l'obiettivo di stimolare il *noticing* e favorire il conseguente sviluppo dell'interlingua.

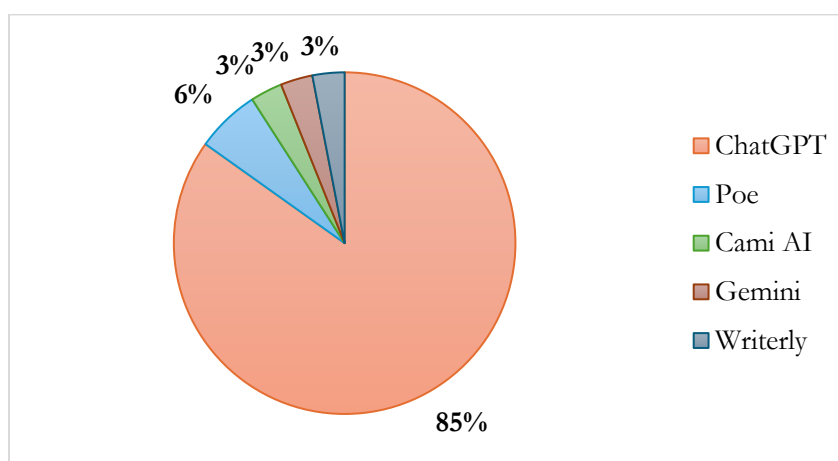
In aggiunta ai quadri teorici più frequentemente adottati, emergono due ulteriori orientamenti nella letteratura analizzata. Da un lato, possiamo osservare una crescente attenzione verso la dimensione affettiva dell'apprendimento linguistico, richiamata da modelli di riferimento quali il *L2 Motivational Self System* (Dörnyei, 2009) e la *Self-Regulated Learning Theory* (Zimmermann, 2000). Gli studi che adottano questo contesto teorico esplorano l'impatto che l'utilizzo dei *chatbot* basati su LLM a supporto dei processi di scrittura in L2/LS determina sulla motivazione e sull'autonomia di apprendimento. In particolare, il feedback personalizzato e interattivo proposto dal *chatbot* è considerato come un'opportunità per promuovere strategie metacognitive attraverso la pianificazione e il monitoraggio costante del proprio processo di scrittura. L'apprendimento è pertanto favorito dall'osservazione, dalla riflessione e dall'adattamento progressivo della propria produzione scritta in risposta al feedback ricevuto.

Dall'altro lato, si rileva un interesse particolare verso l'adozione di *framework* concettuali a supporto di un'integrazione efficace della IAG nei contesti edulinguistici. In questa direzione, modelli di riferimento come il TPACK (*Technological Pedagogical Content Knowledge*) di Koehler *et al.* (2013) e il SAMR (*Substitution, Augmentation, Modification, Refefinition*) di Puentedura (2014) offrono strumenti teorici utili per promuovere un uso pedagogicamente consapevole della IAG a supporto delle abilità di scrittura in L2/LS. Allo stesso modo, emerge la consapevolezza di sviluppare nuove forme di *AI literacy* (Warschauer *et al.*, 2023) che conducano allo sviluppo di competenze trasversali, quali la capacità di interagire efficacemente con i modelli linguistici generativi attraverso strategie di *prompting*, finalizzate alla valorizzazione del potenziale conversazionale della IAG mantenendo al contempo un approccio *human-centered* all'interno del processo di apprendimento linguistico.

3.2. Tipologie di LAG e strategie di utilizzo

L'integrazione della IAG per assistere la produzione scritta nei percorsi di apprendimento linguistico in L2/LS si realizza attraverso scelte eterogenee, ma con una chiara tendenza all'adozione di ChatGPT come tecnologia dominante, a strategie di *prompting* prevalentemente fornite dal docente e utilizzate all'interno di contesti didattici tradizionali. I dati graficamente riportati nella Figura 8 evidenziano una netta convergenza verso l'utilizzo di ChatGPT (85%), che si configura come la tecnologia di riferimento nella quasi totalità degli studi esaminati. Solo una parte percentualmente ridotta degli articoli fa riferimento ad altre tipologie di strumenti come Poe (6%), mentre Gemini, CamiAI e Writerly sono utilizzati nel 3% delle ricerche. L'adozione quasi esclusiva di ChatGPT è interpretabile come il risultato della sua accessibilità, visibilità pubblica e facilità d'uso, costituendo le premesse per la costruzione di evidenze cumulative. Tuttavia, la scarsa presenza di studi che istituiscono confronti sistematici tra *chatbot* alternativi potrebbe costituire un limite per la replicabilità e la validità comparativa degli studi.

Figura 8. Distribuzione degli studi in base alla tipologia di LAG adottata



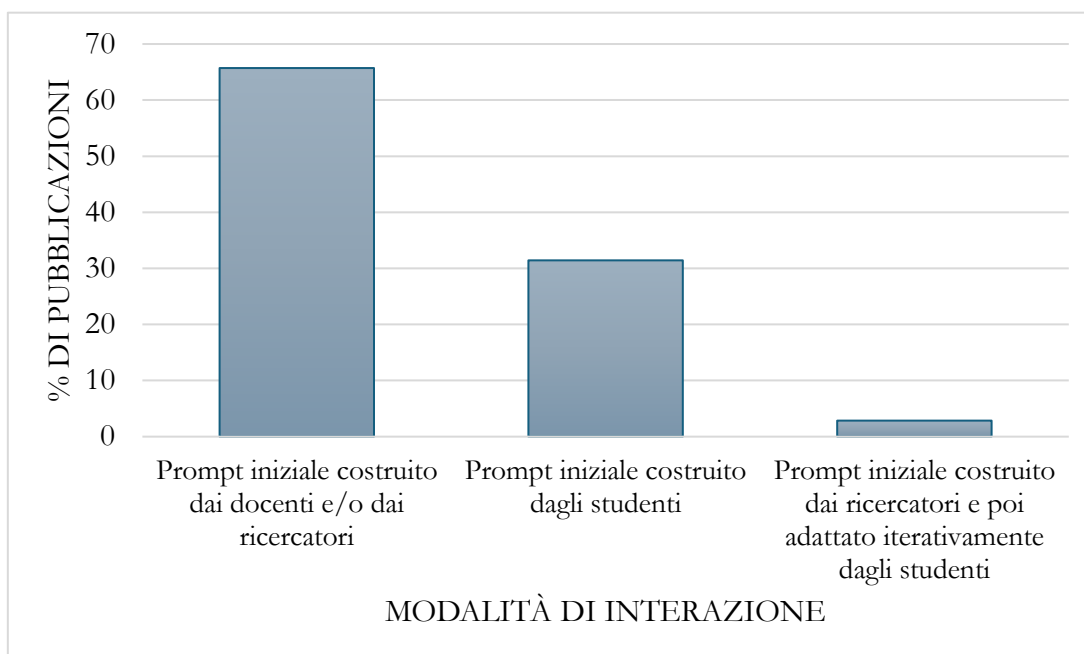
3.2.1. Modalità di interazione con la LAG e contesti di utilizzo

L'implementazione delle applicazioni basate su IAG nella didattica della scrittura in L2/LS si realizza attraverso varie modalità di interazione e all'interno di contesti educativi

diversi, che spaziano da ambienti formali come la classe e il laboratorio didattico, a modalità più informali e autonome quali l'interazione asincrona da parte dell'apprendente al di fuori del contesto accademico, fino a utilizzi ibridi che integrano momenti di lavoro guidato e attività svolte in autonomia. L'articolazione di tali dimensioni determina in misura significativa le modalità con cui la IAG è introdotta nei percorsi didattici, influenzando la tipologia e la qualità dell'esperienza d'uso. In questa prospettiva, le modalità di generazione e somministrazione dei *prompt* costituiscono una dimensione significativa nella progettazione didattica, poiché definiscono la qualità dell'interazione tra l'apprendente e lo strumento generativo, la natura dell'input linguistico e il grado di supporto fornito all'apprendente attraverso il processo di *scaffolding* durante le diverse fasi della produzione scritta.

Le modalità di interazione con gli strumenti basati sulla IAG mostrano una tendenza generale verso strategie di *prompting* standardizzate, che suggeriscono la volontà di mantenere un maggior controllo durante l'intervento didattico e garantire un utilizzo etico dello strumento. Nello specifico, la Figura 9 mostra come quasi i due terzi degli studi analizzati prevedono che il *prompt* sia fornito dai ricercatori o dai docenti (66%), mentre solo una minoranza prevede che gli studenti generino autonomamente i propri input testuali (31%) o, dato un *prompt* iniziale precostituito, possano modificarlo a piacimento (3%). Se da un lato questa impostazione metodologica potrebbe contribuire a una maggiore comparabilità dei risultati tra i partecipanti, dall'altro potrebbe incentivare un uso passivo e meramente esecutivo da parte degli apprendenti. Al contrario, offrire la possibilità di personalizzare l'interazione con *chatbot* basati su LLM durante il processo di scrittura potrebbe favorire un maggiore sviluppo delle competenze metacognitive e dell'autonomia di apprendimento.

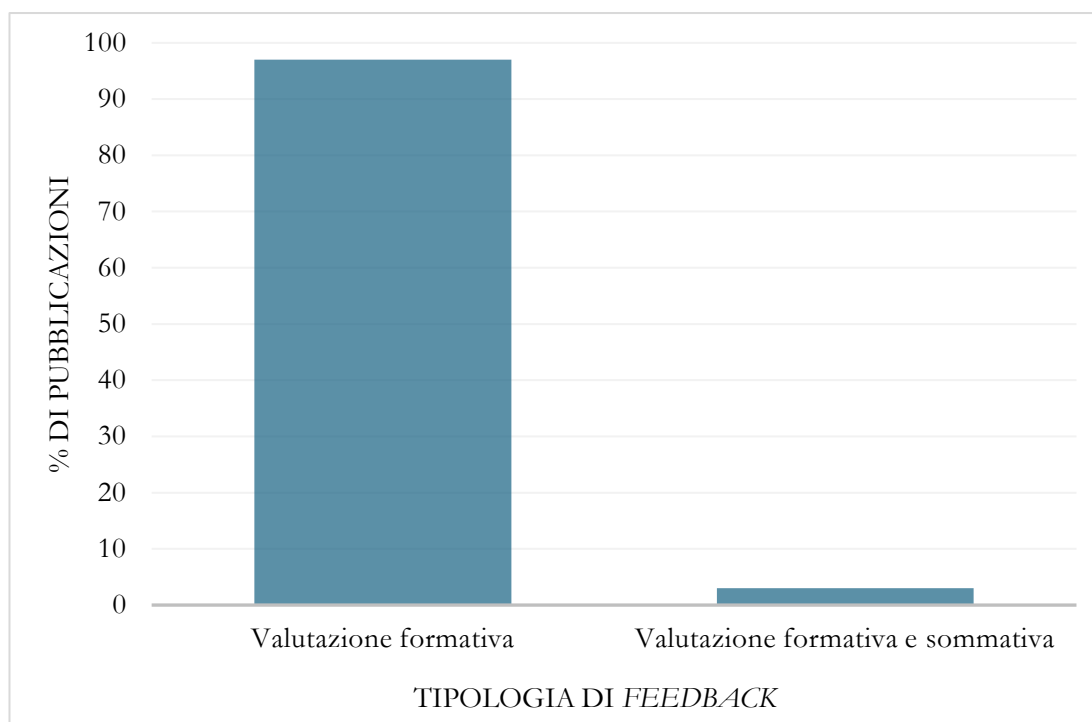
Figura 9. Strategie di *prompting* adottati durante l'interazione con la IAG



Anche la tipologia di feedback fornito rappresenta un elemento importante nel configurare il rapporto tra l'apprendente e l'applicazione basata su IAG all'interno dell'esperienza formativa. La Figura 10 mostra che soltanto il 3% dei contributi prevede una fase conclusiva di valutazione sommativa, strutturata attraverso rubriche analitiche

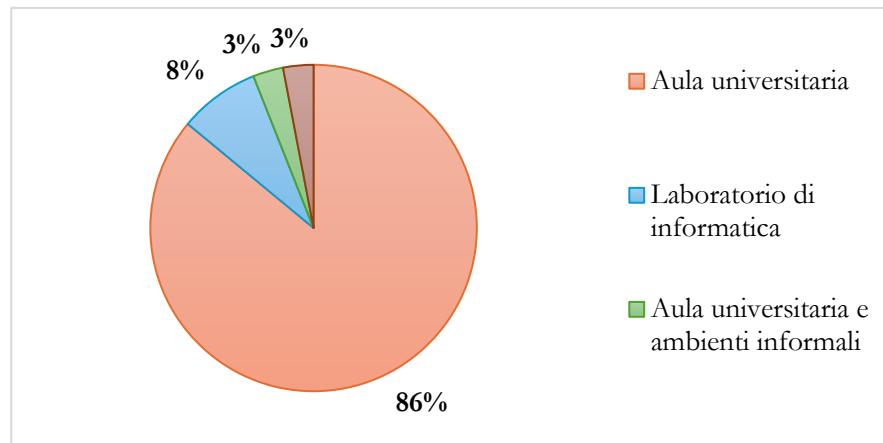
applicate al testo prodotto per misurare l'impatto che l'utilizzo del *chatbot* ha avuto sulla qualità della produzione scritta (a scopo esemplificativo, cfr. Lytovchenko *et al.*, 2025). Nella quasi totalità degli studi (97%) il feedback fornito durante tutte le fasi di scrittura in L2/LS riflette la natura formativa della valutazione, enfatizzandone la «funzione di monitoraggio del processo d'apprendimento e del percorso d'insegnamento [...] che prevede un maggiore intervento autovalutativo da parte dello studente, che consente di modificare nei tempi e nei modi, riprendere e riproporre un determinato input didattico» (Mezzadri, 2024: 294).

Figura 10. *Tipologia del feedback in funzione della valutazione delle abilità di scrittura*



3.2.2. Contesti di apprendimento linguistico

Per quanto riguarda il contesto educativo in cui sono stati condotti gli studi esaminati, possiamo osservare come nel 97% dei casi l'implementazione degli strumenti basati sulla IAG avviene in ambienti di apprendimento formali. In particolare, lo spazio didattico più rappresentato è quello universitario tradizionale, durante le lezioni in aula (86%), seguito dai laboratori di informatica (8%) e da una combinazione di entrambi (3%). Soltanto il 3% prevede invece momenti di utilizzo in contesti ibridi, che combinano contesti formali e informali, come illustrato nella Figura 11. La centralità degli ambienti formali è interpretabile come scelta funzionale al monitoraggio e alla supervisione intenzionale dell'interazione tra *chatbot* basati su LLM e apprendente. Tuttavia, sebbene in casi isolati, la natura stessa della tecnologia, che ne permetterebbe l'utilizzo asincrono, interattivo e personalizzato, potrebbe costituire un'occasione per promuovere una maggiore autonomia e flessibilità nei processi di apprendimento.

Figura 11. *Contesti di apprendimento formali e informali in relazione all'uso della LAG*

3.3. Ruolo educativo della LAG per le abilità di scrittura

L'analisi degli studi restituisce un quadro articolato e multifunzionale del ruolo educativo svolto dalle applicazioni basate sulla IAG per l'apprendimento linguistico in L2/LS. La natura polifunzionale dei *chatbot* fa sì che il numero complessivo delle occorrenze superi quello degli studi analizzati, poiché ogni contributo può descrivere diversi utilizzi del *chatbot* che ricadono in più ambiti educativi contemporaneamente. È stata perciò adottata una codifica tematica multipla che consente di attribuire a ciascuno studio più temi, qualora emergessero più ruoli e/o funzioni pedagogiche dello strumento. Di conseguenza, il numero totale delle occorrenze tematiche supera il numero degli studi analizzati, riflettendo la natura polifunzionale delle tecnologie basate sulla IAG.

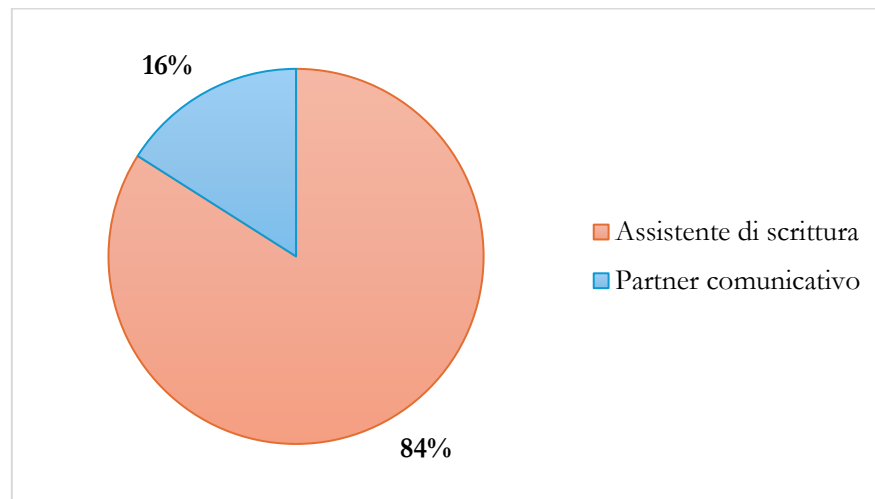
Questa scelta richiede tuttavia alcune fondamentali riflessioni metodologiche: da un lato, uno sbilanciamento quantitativo tra le categorie considerate; dall'altro, il rischio di una sovrainterpretazione dei ruoli attribuiti ai *chatbot* nei diversi studi. Ciononostante, riteniamo che l'approccio adottato possa maggiormente riflettere la natura multifunzionale di questi strumenti, che non svolgono quasi mai un ruolo educativo univoco, ma si adattano a contesti, obiettivi e interazioni diversi, intervenendo su più abilità e utilizzando metodologie ibride. Limitarsi a una codifica esclusiva semplificherebbe la realtà descritta dagli studi, rendendo più difficile cogliere la coesistenza di ruoli diversi e l'intrinseca complessità strutturale degli interventi.

Tra i ruoli educativi attribuiti alle applicazioni basate su IAG nei contributi esaminati, prevale quello di assistente alla scrittura, che ricorre nell'84% dei casi (Figura 12). In tale configurazione, il *chatbot* agisce come mediatore nel processo compositivo, intervenendo trasversalmente dalla pianificazione alla revisione del testo. Questo ruolo assume declinazioni funzionali diverse, ma complementari, che cooperano all'interno dell'ambiente conversazionale per sostenere l'apprendente in modo continuo e adattivo: più della metà delle occorrenze (58%) fa riferimento alla funzione di correttore grammaticale e lessicale, a cui si affiancano, in misura minore, il supporto all'organizzazione testuale (3%) e il contributo allo *scaffolding* generativo delle idee (23%), suggerendo un utilizzo che si estende oltre la revisione normativa. Parallelamente, nel 16% le applicazioni basate su IAG adottano il ruolo di partner comunicativi, assumendo le funzioni di interlocutore dialogico con cui l'apprendente può negoziare il significato e sviluppare consapevolezza metalinguistica. In questo senso, il *chatbot* non si limita a intervenire su sollecitazione dell'apprendente, ma può attivare dinamiche dialogiche ed

esplorative che favoriscono anche lo sviluppo del pensiero critico e delle capacità di *problem solving*.

L'analisi delle modalità di utilizzo mostra inoltre che il ruolo di partner comunicativo emerge sempre in associazione a quello di assistente di scrittura. Ciò suggerisce come la dimensione dialogica dei *chatbot* basati su LLM non costituisca una funzione autonoma, ma si integri e si sviluppi all'interno del quadro di assistenza alla produzione testuale, nel quale esso diventa interlocutore dialogico e sostegno al processo di generazione e revisione del testo. La compresenza dei ruoli che le applicazioni supportate dai LLM svolgono riflette dunque la natura stratificata della scrittura, intesa non solo come abilità linguistica, ma come competenza metacognitiva che richiede sostegno su più livelli: dalla microstruttura formale all'organizzazione delle idee, fino alla riflessione sui propri processi cognitivi e linguistici.

Figura 12. Ruoli educativi attribuiti alla LAG durante il processo di scrittura

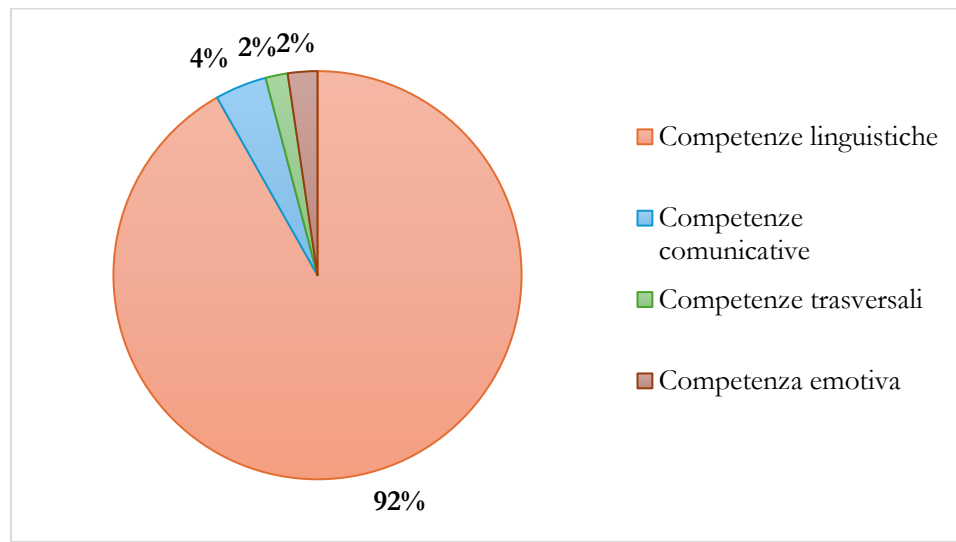


Dalla sintesi dei risultati emerge inoltre come l'utilizzo di applicazioni supportate da IAG solleciti una molteplicità di aspetti legati alla produzione scritta (Figura 13). L'analisi delle occorrenze mostra una netta prevalenza di interventi riferibili alle competenze linguistiche, che rappresentano il 92% dei casi e comprendono sia elementi formali del linguaggio (44%) sia di contenuto (48%). Questa configurazione riflette un utilizzo dello strumento che si colloca in continuità con approcci glottodidattici di tipo tradizionale in cui il *chatbot* viene utilizzato secondo logiche simili a quelle dei sistemi AWE, con un approccio normativo e incentrato sulla revisione, la correzione e l'adeguamento alle norme linguistiche. Si rileva, tuttavia, un'attenzione crescente verso lo sviluppo delle competenze comunicative (4%), che costituiscono un polo emergente orientato a promuovere l'efficacia comunicativa, l'appropriatezza verso la consegna e le convenzioni accademiche, indicando un uso più situato e consapevole della scrittura nel contesto d'apprendimento.

La distribuzione dei dati mostra inoltre una tendenza predominante all'utilizzo dell'IA orientato al supporto linguistico-formale, ma al contempo suggerisce un progressivo ampliamento delle sue funzioni in direzione di una prospettiva olistica, che riconosce le dimensioni cognitive, affettive e sociali del processo di scrittura. Sebbene la maggior parte degli studi si concentri sulla valutazione dell'accuratezza linguistica, emerge tuttavia un interesse, seppur limitato, in riferimento allo sviluppo delle competenze trasversali (2%), come la creatività e la collaborazione tra pari, e alle competenze emotive (2%) degli

apprendenti, quali la gestione dell'ansia, della motivazione e del senso di autoefficacia. Questi dati indicano un impiego dei chatbot basati sui LLM che, pur non abbandonando le funzioni tradizionali di supporto linguistico, si orienta verso un approccio più complesso e integrato alla scrittura in L2/LS, in cui le competenze linguistiche vengono affiancate alle competenze trasversali e affettive, in linea con una visione didattica più ampia e orientata allo sviluppo globale delle competenze dell'apprendente.

Figura 13. *Competenze legate alle abilità di scrittura durante l'uso della IAG*



4. DISCUSSIONE

Il rapido sviluppo tecnologico nel campo dell'IA, segnato dall'avvento della IAG, si riflette marcatamente anche in ambito edulinguistico. Un segnale tangibile di questo fenomeno è rappresentato dalla crescita esponenziale della produzione scientifica di riferimento, che negli ultimi due anni ha decuplicato il numero di pubblicazioni riguardanti l'utilizzo delle tecnologie basate sulla IAG per il potenziamento delle abilità di scrittura in L2/LS, confermando il forte interesse della comunità scientifica verso questo nuovo ambito di ricerca. Tuttavia, la distribuzione geografica evidenzia un marcato sbilanciamento a favore del continente asiatico, con una prevalenza di contributi provenienti da Cina, Thailandia e Indonesia, a fronte di una significativa sottorappresentazione dei contesti africani, americani, oceaniani ed europei. Questo dato trova riscontro anche nelle recenti revisioni sistematiche della letteratura sull'AIED per l'apprendimento linguistico (cfr. Crompton *et al.*, 2024; Zhu, Wang, 2025) e potrebbe essere interpretato alla luce della forte pressione istituzionale sull'apprendimento dell'inglese come lingua veicolare (Kirkpatrick, 2012), che ha favorito l'adozione di soluzioni didattiche percepite come efficienti e scalabili, tra cui l'impiego di LLM. Tale polarizzazione, espressione dell'intensa attività di ricerca sviluppata in area asiatica, contribuisce a delineare un panorama conoscitivo parziale. La concentrazione di studi provenienti da contesti relativamente omogenei produce infatti un bias geografico che limita le opportunità di generalizzabilità e di trasferibilità dei risultati a sistemi educativi e socioculturali diversi. Tali considerazioni trovano un parallelo anche nella scelta del

disegno di ricerca, che rappresenta una dimensione rilevante per comprendere il panorama degli studi sulle applicazioni della IAG a supporto della scrittura in L2/LS.

L'analisi condotta rivela una marcata inclinazione verso disegni di tipo quasi-sperimentale e sperimentale, spesso associati a un impiego prevalente di approcci metodologici a metodi misti e con un campione di partecipanti di medio-piccole dimensioni. Questi risultati ci permettono di avanzare alcune osservazioni. In primo luogo, la scelta di combinare strumenti di analisi qualitativa e quantitativa, riscontrabile anche in Zhu e Wang (2024), riflette l'esigenza di adottare una prospettiva olistica per cogliere la complessità del processo di scrittura mediata da LLM, attraverso la raccolta integrata di dati quantitativi relativi alle competenze linguistiche con evidenze qualitative riguardanti le percezioni e le strategie metacognitive adottate dagli apprendenti. In secondo luogo, la prevalenza di approcci quasi-sperimentali, pur evidenziando la volontà di misurare l'impatto dell'uso della IAG sulla produzione scritta in L2/LS, comporta una limitazione nella comparabilità dei risultati e una difficile attribuzione degli esiti osservati all'intervento stesso. A ciò si aggiunge la scarsità di studi longitudinali, che riduce la possibilità di monitorare l'effettiva evoluzione delle abilità di scrittura oltre l'immediata interazione con lo strumento (cfr. Zawacki-Richter *et al.*, 2019).

Allo stesso tempo, emerge una crescente attenzione ad implementare l'utilizzo di *chatbot* basati su LLM all'interno di *framework* di riferimento in ambito educativo. Dall'analisi dei risultati si conferma infatti il ruolo cruciale della scrittura in L2/LS come abilità complessa che coinvolge competenze linguistiche, strategiche ed emotive e che si realizza in contesti comunicativi autentici. In questa prospettiva, la capacità dei *chatbot* di fornire un supporto immediato, personalizzato e interattivo durante tutte le fasi del processo di scrittura si colloca all'interno di un approccio costruttivista, secondo cui l'interazione con lo strumento consente agli apprendenti di affrontare attività linguisticamente complesse che ricadono nella loro ZPD (Vygotsky, 1978). Tali risultati sono coerenti con quanto riportato dall'analisi condotta da Chen *et al.* (2020a) sull'evoluzione della ricerca sull'AIED nell'ultimo ventennio, che evidenzia una predominanza dell'approccio costruttivista già prima della diffusione della IAG nei contesti educativi.

Un'ulteriore dimensione emergente è rappresentata dall'adozione di *framework* per orientare l'integrazione degli strumenti basati sulla IAG nelle abilità di scrittura in L2/LS da una prospettiva educativa. Tra questi, il modello SAMR (Puentedura, 2014) e il *framework* TPACK (Koehler *et al.*, 2013) emergono come riferimenti per interpretare il ruolo delle tecnologie nei processi di apprendimento e insegnamento. Ad esempio, lo studio di Muslimin *et al.* (2024) adotta il SAMR per analizzare l'impatto del feedback generato da CamiAI. La progressiva integrazione del *chatbot* attraverso i quattro livelli (cfr. § 3.1.3) ha evidenziato non solo un miglioramento della qualità della produzione degli apprendenti, ma anche un coinvolgimento più attivo, critico e riflessivo. Analogamente, Tseng e Lin (2024) utilizzano il TPACK per analizzare l'impatto di ChatGPT sulla scrittura in L2/LS, mostrando come una progettazione didattica che integra coerentemente competenze tecnologiche, pedagogiche e disciplinari aumenti la qualità delle produzioni scritte e promuova un uso più consapevole della IAG, attraverso la formulazione di *prompt* efficaci e la valutazione critica dell'interazione. Quest'ultimo aspetto è esplorato anche da Kurt e Kurt (2024), che sottolineano l'importanza di sviluppare una solida *AI literacy* negli apprendenti.

Sebbene questi contributi rappresentino esempi significativi di integrazione tra *framework* concettuali e pratica didattica, il panorama complessivo della ricerca risulta, nel 45% dei casi, caratterizzato da una mancata esplicitazione dei riferimenti teorici impiegati. Questo dato, già evidenziato in recenti revisioni sistematiche sull'AIED (cfr. Chen *et al.*, 2020a, 2020b; Ouyang *et al.*, 2022), sembrerebbe riflettere la natura fortemente

interdisciplinare del settore, che tende a privilegiare l'attenzione verso gli aspetti di sviluppo e di prestazione tecnica degli strumenti. Tuttavia, questo panorama pone questioni rilevanti per la ricerca futura poiché potrebbe, da un lato, spostare il focus dall'analisi delle implicazioni glottodidattiche verso l'innovazione tecnologica e, dall'altro, compromettere la possibilità di sviluppare in futuro evidenze trasferibili e comparabili sul ruolo e sulle potenzialità dei chatbot basati su LLM nel supporto alle abilità di scrittura in L2/LS. In tale prospettiva, integrare l'utilizzo degli applicativi basati sui LLM all'interno di *framework* concettuali in ambito educativo rappresenta una condizione fondamentale per promuovere un uso formativo e critico della IAG nella glottodidattica.

In questa direzione sembra orientarsi anche l'analisi dei dati relativi alle modalità di interazione tra gli apprendenti e i *chatbot*. Nonostante la tendenza predominante all'uso di *prompt* precostituiti da parte di docenti o ricercatori, alcuni studi coinvolgono attivamente gli studenti nella generazione autonoma dei *prompt* (Muslimin *et al.*, 2024; Tsai *et al.*, 2024). Questa scelta metodologica si associa ad una maggiore centralità dello studente come soggetto attivo nei processi di apprendimento e scrittura, promuovendo pratiche di interazione più interattive e contestualizzate. L'elaborazione autonoma dei *prompt* potrebbe favorire, per esempio, lo sviluppo del pensiero critico, poiché gli apprendenti sono chiamati a valutare l'adeguatezza delle risposte dei *chatbot*, a riflettere sui propri obiettivi comunicativi e a riformulare le richieste in maniera iterativa. In una prospettiva socioculturale, l'acquisizione di competenze legate al *prompting*, all'interno di una didattica orientata allo sviluppo dell'autonomia, rappresenta non soltanto uno strumento operativo per interagire efficacemente con le tecnologie basate sulla IAG, bensì un requisito fondamentale all'interno di percorsi di alfabetizzazione critica all'IA, al fine di promuovere un uso consapevole e responsabile nei contesti educativi.

Anche il ruolo educativo dei *chatbot* basati su LLM nei processi di apprendimento è ampiamente riconosciuto, sebbene declinato in modalità diverse a seconda del contesto didattico. Numerosi studi ne evidenziano l'uso come assistenti di scrittura (tra i contributi si vedano Strobl *et al.*, 2024; Chen, Gong, 2025), ma emergono anche funzioni più complesse, come il supporto allo sviluppo di idee e la facilitazione del pensiero critico (Werdingingsih *et al.*, 2024a; Tseng, Lin, 2024). Altri studi, come quello condotto da Niloy *et al.* (2024), adottano una prospettiva critica, mettendo in guardia sugli effetti potenzialmente negativi della IAG sulla produzione scritta, tra cui la riduzione della creatività e dell'originalità dei contenuti. Questa polarizzazione nei risultati evidenzia la duplice natura dell'interazione con i *chatbot*. Se da un lato possono infatti stimolare le competenze linguistiche, metacognitive e trasversali, dall'altro ne enfatizzano i rischi connessi ad un utilizzo acritico.

Questa eterogeneità di prospettive si riflette anche nella capacità degli apprendenti di interpretare, valutare e applicare criticamente il feedback ricevuto dai *chatbot* basati su LLM. Sebbene nella quasi totalità degli studi analizzati la valutazione ha finalità formative e la maggioranza di essi converge nel ritenere il feedback necessario per consolidare le abilità di scrittura in L2/LS, emergono delle discrepanze nella modalità con cui esso viene percepito e rielaborato dai discenti. La ricerca condotta da Seo (2024) mostra come un uso passivo della tecnologia possa ridurre l'efficacia del feedback; in particolare, l'attenzione degli studenti si è concentrata quasi esclusivamente sulle correzioni formali, con una scarsa rielaborazione concettuale e un coinvolgimento attivo limitato. Similmente, Wang (2024) enfatizza che in assenza di una mediazione guidata con lo strumento, il feedback rischia di essere percepito come superficiale e ridondante, ostacolandone il valore formativo.

Al contrario, un'interazione consapevole e attiva con queste tecnologie durante l'attività di scrittura può favorire lo sviluppo della *feedback literacy*, un'area emergente che esplora la capacità degli apprendenti di riconoscere e apprezzare il valore della valutazione

formativa, autoregolare le proprie risposte emotive, decodificare criticamente il feedback ricevuto e tradurlo in azioni efficaci di miglioramento (Molloy *et al.*, 2020). Bai e Wei (2024) mostrano come l'interazione con ChatGPT possa attivare processi metacognitivi, spronando gli studenti a riflettere e rielaborare criticamente sulla propria produzione scritta. Analogamente, gli studi condotti da Bacon e Maneerutt (2024) e Oktarin *et al.* (2024) indicano come l'interazione consapevole con il *chatbot* possa rafforzare lo sviluppo di un'alfabetizzazione critica al feedback, contribuendo allo stesso tempo allo sviluppo delle competenze linguistiche in L2/LS. Questi risultati suggeriscono come il feedback possa trasformarsi da semplice correzione automatica degli errori a strumento formativo che stimola l'autonomia di apprendimento, se integrato all'interno di pratiche didattiche che promuovono un'interazione critica, consapevole e formativa con la IAG.

L'analisi delle abilità linguistiche per cui gli apprendenti richiedono più frequentemente supporto ai *chatbot* rivela una tendenza a replicare caratteristiche tipicamente attribuite agli strumenti di AWE (cfr. § 1.2). Nonostante il potenziale interattivo e adattivo dei modelli linguistici generativi, molti studi mostrano come il loro utilizzo sia ancora prevalentemente orientato verso feedback mirati alla correzione degli elementi formali della lingua, in particolare alla correttezza grammaticale e lessicale (Mahapatra, 2024; Wang, 2024). Questa modalità di interazione, in continuità con quanto osservato da Huang *et al.* (2024), rischia di limitarne l'efficacia educativa, soprattutto a scapito dello sviluppo di strategie metacognitive, sociali e affettive durante il processo di scrittura. In tal senso, sebbene alcuni studi si pongano in controtendenza, sottolineandone i benefici nel migliorare l'organizzazione testuale e la rielaborazione dei contenuti (Bacon, Maneerutt, 2024; Werdiningsih *et al.*, 2024b), il potenziale trasformativo della IAG risulta ancora parzialmente inespresso. Ciò suggerisce pertanto la necessità di un'implementazione didattica più consapevole, che ne promuova un utilizzo critico e polifunzionale.

5. CONCLUSIONE

La presente *scoping review* ha offerto una panoramica sistematica della letteratura scientifica sull'utilizzo dell'IAG a supporto delle abilità di scrittura in L2/LS in contesti accademici. I risultati emersi evidenziano il carattere ancora emergente di questo ambito di ricerca.

L'evoluzione temporale entro cui si collocano la maggior parte degli studi conferma la rapidità con cui la IAG è stata integrata nei contesti educativi. Inoltre, l'asimmetria geografica della letteratura verso il continente asiatico riflette un panorama parziale della ricerca, che potrebbe non riflettere adeguatamente gli utilizzi delle tecnologie IAG in contesti educativi pluriculturali e plurilingui.

La natura emergente del fenomeno sembra aver influito sull'assenza di studi longitudinali in grado di misurare gli effetti a medio e lungo termine degli strumenti IAG sull'apprendimento linguistico. Parallelamente, la ridotta rappresentatività dei campioni impiegati, prevalentemente di dimensioni ridotte e circoscritti ai singoli corsi universitari, limita la possibilità di estendere le evidenze a contesti più ampi e differenziati, sollevando interrogativi sulla validità delle evidenze ottenute, soprattutto rispetto a popolazioni di apprendenti più eterogenee per livello linguistico, background culturale e bisogni formativi.

Inoltre, il ruolo pedagogico dei *chatbot* viene frequentemente declinato in chiave assistenzialistica e correttiva, con una preponderanza di funzioni legate alla correttezza linguistica e alla revisione. L'interesse principale degli studi rimane concentrato sul miglioramento delle competenze linguistiche formali, marginalizzando le riflessioni

relative alle competenze comunicative, trasversali e affettive e i rischi concreti di dipendenza dallo strumento, standardizzazione linguistica e riduzione dell'autonomia autoriale.

Infine, l'inquadramento teorico che caratterizza buona parte della letteratura analizzata evidenzia la necessità di una maggiore interdisciplinarietà. Quasi la metà degli studi non adotta un *framework* concettuale esplicito, in linea con la tendenza della letteratura relativa all'AIED a privilegiare le dimensioni funzionali e legate alla performance dell'uso dei *chatbot* basati su LLM, a scapito di implicazioni glottodidattiche e pedagogiche. Ciò limita la possibilità di collocare l'uso di applicazioni basati su questi strumenti entro quadri teorici coerenti con i principi pedagogici, rendendo più difficile la costruzione di cornici interpretative condivise.

5.1. Limiti della *scoping review*

Il processo di revisione sistematica adottato nella *scoping review* non è esente da limitazioni di natura metodologica. Innanzitutto, la scelta di selezionare articoli redatti esclusivamente in lingua inglese ha comportato l'esclusione di contributi potenzialmente rilevanti, ma pubblicati in altre lingue. Sebbene la maggior parte degli studi presenti nelle principali banche dati adottino la lingua inglese, questa scelta metodologica potrebbe aver limitato l'accesso a prospettive glottodidattiche significative per l'analisi dell'utilizzo di strumenti basati sulla IAG per le abilità di scrittura in L2/LS. Un limite simile riguarda inoltre la scelta delle banche dati e la stringa di ricerca utilizzate per l'identificazione degli studi. Nonostante i ricercatori abbiano condotto il processo di interrogazione in Scopus, EBSCOHost ed ERIC in maniera sistematica al fine di garantire una ricerca il più esaustiva e coerente possibile della letteratura scientifica di riferimento, è plausibile che alcune pubblicazioni pertinenti, presenti in archivi meno noti o indicizzate con termini diversi, possano essere state escluse dal corpus dei contributi analizzati.

5.2. Sviluppi futuri

Alla luce dei risultati emersi sull'uso delle applicazioni basate sulla IAG nella scrittura in L2/LS in contesti accademici, questa *scoping review* offre un quadro organico utile per orientare la ricerca futura, costituendo un punto di partenza per l'elaborazione di approcci didattici informati e lo sviluppo di nuove domande di ricerca.

Una prima direzione riguarda la relazione tra l'autorialità testuale, l'autonomia linguistica e la costruzione dell'*Ideal L2 Writing Self*. Se l'IAG sembra offrire un supporto nel ridurre l'ansia, fornire modelli di scrittura e migliorare la leggibilità del testo, dall'altro lato può anche generare effetti di deresponsabilizzazione, facendo sentire gli studenti editori di testi pre-generati. Sarebbe perciò necessario indagare come gli studenti percepiscano la paternità del testo prodotto con il supporto di *chatbot* basati su LLM e in che modo tale percezione incida sulla motivazione, sulla *self-efficacy* e sull'autonomia linguistica.

Particolare attenzione merita inoltre il tema delle competenze di *prompting*, attualmente considerato per lo più come variabile sperimentale o tecnica. Le ricerche future potrebbero esplorare criticamente l'interazione uomo-IA indagando non solo le modalità con cui gli studenti formulano il *prompt*, ma anche il livello di consapevolezza linguistica, strategica e metacognitiva che accompagna tale processo. In quest'ottica, potrebbe essere interessante promuovere percorsi formativi orientati allo sviluppo di competenze

trasversali di *prompting*, quali la capacità di pensiero critico, per un'interazione efficace e funzionale con la IAG.

Si ritiene inoltre auspicabile investigare come integrare in modo equilibrato il feedback umano e il feedback generato dall'IA. Alcuni studi (Escalante *et al.*, 2023; Bacon, Maneerutt, 2024) confermano che interventi ibridi producono effetti significativi in termini di *engagement*, consapevolezza metalinguistica e qualità testuale. Ciò rende necessario rilevare non soltanto le percezioni degli insegnanti in merito all'efficacia o all'usabilità degli strumenti basati su LLM, ma indagare anche il loro ruolo all'interno della mediazione tra strumento e apprendimento linguistico. In questa prospettiva, risulta utile fare riferimento a modelli teorici in grado di restituire la complessità del ruolo docente nella progettazione didattica assistita da IA. Modelli come TPACK (Koehler *et al.*, 2013) rappresentano un riferimento operativo importante per garantire un'integrazione pedagogicamente fondata e orientata dell'IAG, che consente di concettualizzare tale ruolo come intersezione dinamica tra conoscenze disciplinari, glottodidattiche e tecnologiche.

Un'ulteriore dimensione di sviluppo futuro potrebbe esplorare in che modo l'uso della IAG possa essere integrato in metodologie didattiche orientate allo sviluppo delle competenze comunicative, superando un approccio normativo prevalentemente orientato alla correttezza grammaticale. In particolare, riteniamo utile esplorare se e in che misura queste tecnologie siano in grado di supportare l'apprendimento su una più ampia varietà di generi testuali, al fine di promuovere un uso autentico e contestualizzato della lingua, una maggiore consapevolezza retorica e competenze comunicative richieste dall'attività di scrittura. Affinché le potenzialità formative della IAG siano pienamente valorizzate, è pertanto necessario che la sua implementazione sia guidata da pratiche didattiche che stimolino processi di mediazione dei significati, pensiero critico e riflessione metacognitiva.

Infine, consideriamo imprescindibile esplorare le potenzialità educative della IAG da una prospettiva etica. Nello specifico, è necessario esplorare le implicazioni del divario digitale, le disuguaglianze nell'accesso e nell'uso delle tecnologie, nonché le modalità e gli esiti dell'interazione dei *chatbot* in contesti socioculturali e linguistici differenti. Riteniamo, in modo convinto, che questa analisi sia centrale per evitare che stereotipi e *bias*, insiti nella natura dei modelli linguistici generativi, contribuiscano ad amplificare forme preesistenti di esclusione, marginalizzazione e standardizzazione linguistica e culturale, garantendo invece un uso equo, consapevole e inclusivo di tali strumenti per l'apprendimento e l'insegnamento delle lingue.

RIFERIMENTI BIBLIOGRAFICI⁴

- Albadarin Y., Saqr M., Pope N., Tukainen M. (2024), "A systematic literature review of empirical research on ChatGPT in education", in *Discover Education*, 3, 60, pp. 1-26: <https://doi.org/10.1007/s44217-024-00138-2>.
- Al-khreshehm M. H. (2024), "The future of artificial intelligence in English language teaching: Pros and cons of ChatGPT implementation through a systematic review", in *Language Teaching Research Quarterly*, 43, 1, pp. 54-80: <https://doi.org/10.32038/ltrq.2024.43.04>.

⁴ I riferimenti bibliografici contrassegnati con (*) fanno parte del corpus finale degli studi inclusi nella *scoping review*.

- Arksey H., O'Malley L. (2005), "Scoping studies: Towards a methodological framework", in *International Journal of Social Research Methodology*, 8, 1, pp. 19-32: <https://doi.org/10.1080/1364557032000119616>.
- *Bacon E. D., Maneerutt G. (2024), "Enhancing English as a Foreign Language academic writing through AI and Peer-assisted learning", in *Journal of Institutional Research South East Asia*, 22, 3, pp. 282-314.
- Balboni P. E. (2006), *The epistemological nature of language teaching methodology*, Guerra, Perugia.
- Bandura A. (1997), *Self Efficacy: The Exercise of Control*, W. H. Freeman, New York.
- *Bai L., Wei Y. (2024), "Exploring EFL learners' integration and perceptions of ChatGPT's text revisions: A three-stage writing task study", in *IEEE Transactions on Learning Technologies*, 17, 1, pp. 2215-2226: <https://doi.org/10.1109/tlt.2024.3491864>.
- *Biju N., Abdelrasheed N. S. G., Bakiyeva K., Prasad K. D. V., Jember B. (2024), "Which one? Ai-assisted language assessment or paper format: an exploration of the impacts on foreign language anxiety, learning attitudes, motivation, and writing performance", in *Language Testing in Asia*, 14, 45, pp. 1-24: <https://doi.org/10.1186/S40468-024-00322-Z>.
- Burstein J., Tetreault J., Chodorow M., Blanchard D., Andreyev S. (2013), "Automated evaluation of discourse coherence quality in essay writing", in Shermis M., Burstein J. (eds.), *Handbook of Automated essay evaluation: Current applications and new directions*, Routledge, New York-London, pp. 267-280.
- Chaka C. (2024), "Reviewing the performance of AI detection tools in differentiating between AI-generated and human-written texts: A literature and integrative hybrid review", in *Journal of Applied Learning and Teaching*, 7, 1, pp. 115-126: <https://doi.org/10.37074/jalt.2024.7.1.14>.
- *Chen C., Gong Y. (2025), "The role of AI-assisted learning in academic writing: A mixed-methods study on Chinese as a Second Language students", in *Education Sciences*, 15, 2, p. 141: <https://Doi.Org/10.3390/Educsci15020141>.
- Chen X., Xie H., Hwang G.-J. (2020a), "A multi-perspective study on artificial intelligence in education: Grants, conferences, journals, software tools, institutions, and researchers", in *Computers and Education: Artificial Intelligence*, 1, 1, pp. 1-11: <https://doi.org/10.1016/j.caeai.2020.100005>.
- Chen X., Xie H., Zou D., Hwang G.-J. (2020b), "Application and theory gaps during the rise of artificial intelligence in education", in *Computers and Education: Artificial Intelligence*, 1, 1, pp. 1-20: <https://doi.org/10.1016/j.caeai.2020.100002>.
- Chiu T. K. F., Moorhouse B. L., Chai C. S., Ismailov M. (2023), "Teacher support and student motivation to learn with artificial intelligence (AI) based chatbot", in *Interactive Learning Environment*, 32, 7, pp. 3240-3256: <https://doi.org/10.1080/10494820.2023.2172044>.
- Cohen L., Manion L., Morrison K. (2017), *Research methods in education* (8th Ed.), Routledge, New York, London.
- Crompton H., Burke D. (2022), "Artificial intelligence in K-12 education", in *SN Social Sciences*, 2, 113, pp. 1-14: <https://doi.org/10.1007/s43545-022-00425-5>.
- Crompton H., Burke D. (2023), "Artificial intelligence in higher education: The state of the field", in *International Journal of Educational Technology in Higher Education*, 20, 22, pp. 1-22: <https://doi.org/10.1186/s41239-023-00392-8>.
- Crompton H., Edmett A., Ichaporia N., Burke D. (2024), "AI and English language teaching: Affordances and challenges", in *British Journal of Educational Technology*, 55, 6, pp. 2503-2529: <https://doi.org/10.1111/bjet.13460>.
- Deane P. (2013), "On the relation between automated essay scoring and modern views of the writing construct", in *Assessing Writing*, 18, 1, pp. 7-24:

<https://doi.org/10.1016/j.asw.2012.10.002>.

- DeKeyser R. (2020), "Skill acquisition theory", in VanPatten B., Keating G. D., Wulff S. (eds.), *Theories in second language acquisition: An introduction* (3rd Ed.), Routledge, New York, Londra, pp. 83-104.
- Daudt H.M.L., van Mossel C., Scott S. J. (2013), "Enhancing the scoping study methodology: A large, inter-professional team's experience with Arksey and O'Malley framework", in *BMC Medical Research Methodology*, 13, 48, pp. 1-9: <https://doi.org/10.1186/1471-2288-13-48>.
- DeKeyser R. (2020), "Skill Acquisition Theory", in VanPatten B., Keating G. D., Wulff S. (eds.), *Theories in second language acquisition: An introduction* (3rd Ed.), Routledge, New York, Londra, pp. 83-104.
- Ding L., Zou D. (2024), "Automated writing evaluation systems: A systematic review of Grammarly, Pigai, and Criterion with a perspective on future directions in the age of generative artificial intelligence", in *Education and Information Technologies*, 29, 1, pp. 14151-14203: <https://doi.org/10.1007/s10639-023-12402-3>.
- Dörnyei Z. (2009), "The L2 motivational self system", in Dörnyei Z., Ushioda E. (eds.), *Motivation, Language Identity and the L2 Self*, Multilingual Matters, Bristol, pp. 9-42.
- *Escalante J., Pack A., Barrett A. (2023), "AI-generated feedback on writing: Insights into efficacy and ENL student preference", in *International Journal of Educational Technology in Higher Education*, 20, 57, pp. 1-20: <https://doi.org/10.1186/S41239-023-00425-2>.
- Hattie J., Timperley H. (2007), The power of feedback, in *Review of Educational Research*, 77, 1, pp. 81-112: <https://doi.org/10.3102/003465430298487>.
- Hedge T. (2011), *Writing*, Oxford University Press, Oxford.
- Holmes W., Luckin R. (2016), *Intelligence unleashed: An argument for AI in education*, Pearson, Ucl Knowledge Lab.
- Huang X., Xu W., Li F., Yu Z. (2024), "A meta-analysis of effects of automated writing evaluation on anxiety, motivation, and second language writing skills", in *The Asia-Pacific Education Researcher*, 33, 6, pp. 1-20: <https://doi.org/10.1007/s40299-024-00865-y>.
- Huang X., Zou D., Cheng G., Xie H. (2023), "Trends, research issues and applications of artificial intelligence in language education", in *Educational Technology and Society*, 26, 1, pp. 112-131: [https://doi.org/10.30191/ETS.202301_26\(1\).0009](https://doi.org/10.30191/ETS.202301_26(1).0009).
- Hyland K., Hyland F. (2006), "Feedback on second language students' writing", in *Language Teaching*, 39, 2, pp. 83-101: <https://doi.org/10.1017/S0261444806003399>.
- Jubier M., Al-Rawe M. F. A., Ghaithi A. A. (2024), "Effect of EditGPT on the learners' autonomy and learning anxiety", in *International Journal of Learning, Teaching and Educational Research*, 23, 8, pp. 369-390: <https://doi.org/10.26803/ijlter.23.8.19>.
- Kasneci E., Sessler K., Küchemann S., Bannert M., Dementieva D., Fischer F., Gasser U., Groh G., Günnemann S., Hüllermeier E., Krusche S., Kutyniok G., Michaeli T., Nerdel C., Pfeffer J., Poquet O., Sailer M., Schmidt A., Seidel T., Stadler M., Weller J., Kuhn J., Kasneci G. (2023), "ChatGPT for good? On opportunities and challenges of large language models for education", in *Learning and Individual Differences*, 103, pp. 1-12: <https://doi.org/10.1016/j.lindif.2023.102274>.
- Katelhön P. (2025), "La rivoluzione silenziosa: IA e mediazione nell'insegnamento delle lingue straniere", in *Italiano LinguaDue*, 17, 1, pp. 475-492: <https://riviste.unimi.it/index.php/promoitals/article/view/29097>.
- Kirkpatrick A. (2012), "English as an Asian lingua franca: The 'Lingua Franca Approach' and implications for language education policy", in *Jelf*, 1, 1, pp. 121-139: <https://doi.org/10.1515/Jelf-2012-0006>.

- *Kurt G., Kurt Y. (2024), "Enhancing L2 writing skills: ChatGPT as an automated feedback tool", in *Journal of Information Technology Education: Research*, 23, 1, pp. 1-24: <https://doi.org/10.28945/5370>.
- Koehler M. J., Mishra P., Cain W. (2013). "What is Technological Pedagogical Content Knowledge (TPACK)?", in *Journal of Education*, 193, 3, pp. 13-19: <https://doi.org/10.1177/002205741319300303>.
- Lai C. Y., Cheung K. Y., Chan C. S. (2023), "Exploring the role of intrinsic motivation in ChatGPT adoption to support active learning: An extension of the technology acceptance model", in *Computers and Education: Artificial Intelligence*, 5, pp. 1-13: <https://doi.org/10.1016/j.cacai.2023.100178>.
- Lee I. (2008), "Understanding teachers' written feedback practices in Hong Kong secondary classrooms", in *Journal of Second Language Writing*, 17, 2, pp. 69-85: <https://doi.org/10.1016/j.jslw.2007.10.001>.
- Levac D., Colquhoun H., O'Brien K. K. (2010), "Scoping studies: Advancing the methodology", in *Implementation Science*, 5, 69, pp. 1-9: <https://doi.org/10.1186/1748-5908-5-69>.
- Li M. (2024), "Leveraging ChatGPT for second language writing feedback and assessment", in *International Journal of Computer-Assisted Language Learning and Teaching*, 14, 1, pp. 1-11: <https://doi.org/10.4018/IJCALLT.360382>.
- Ling Jen S., Salam A. R. B. H. (2024), "Using artificial intelligence for essay writing", in *Arab World English Journal, Special Issue on ChatGPT*, pp. 90-99: <https://dx.doi.org/10.24093/awej/ChatGPT.5>.
- Long M. H. (1981), "Input, interaction, and second-language acquisition", in *Annals of the New York Academy of Sciences*, 379, 1, pp. 259-278: <https://doi.org/10.1111/j.1749-6632.1981.tb42014.x>.
- Long P., Siemens G. (2011), "Penetrating the fog: Analytics in learning and education", in *Educause Review*, 46, 5, pp. 31-40.
- *Lytovchenko I., Lavrysh Y., Synekop O., Lukianenko V., Chugai O., Shastko I. (2025), "The use of ChatGPT in task-based ESP learning at university: Does it make a difference?", in *International Journal of Interactive Mobile Technologies*, 19, 2, pp. 4-22: <https://doi.org/10.3991/ijim.v19i02.51115>.
- *Mahapatra S. (2024), "Impact of Chatgpt on ESL students' academic writing skills: a mixed methods intervention study", in *Smart Learning Environments*, 11, 9, pp. 9-18: <https://doi.org/10.1186/s40561-024-00295-9>.
- Mattioda M. M., Molino A., Cinato L. (2023), "L'intelligenza artificiale per la traduzione: orizzonti, pratiche e percorsi formativi", in *mediAzioni*, 39, 1, pp. 1-17: <https://doi.org/10.6092/issn.1974-4382/18784>.
- Mezzadri M. (2024), *Apprendere, insegnare e valutare le lingue*, Franco Cesati Editore, Firenze.
- Molloy E., Boud D., Henderson M. (2020), "Developing a learning-centred framework for feedback literacy", in *Assessment and Evaluation in Higher Education*, 45, 4, pp. 527-540: <https://doi.org/10.1080/02602938.2019.1667955>.
- Mousavinasab E., Zarifasanaiey N., Kalhor N., Rakhshan M., Keikha M., Saeedi M. G. (2021), "Intelligent tutoring systems: A systematic review of characteristics, applications, and evaluation methods", in *Interactive Learning Environments*, 29, 1, pp. 142-16: <https://doi.org/10.1080/10494820.2018.1558257>.
- *Muslimin A. I., Mukminatien N., Ivone F. M. (2024), "Evaluating CamiAI across SAMR stages: Students' achievement and perceptions in EFL writing instruction", in *Online Learning Journal*, 28, 2, pp. 1-19: <https://doi.org/10.24059/Olj.V28i2.4246>.
- *Niloy A. C., Akter S., Sultana N., Sultana J., Rahman S. I. U. (2024), "Is ChatGPT a menace for creative writing ability? An experiment", in *Journal of Computer Assisted Learning*, 40, 2, pp. 919-930: <https://doi.org/10.1111/Jcal.12929>.

- *Oktarin I. B., Saputri M. E. E., Magdalena B., Hastomo T., Maximilian A. (2024), "Leveraging ChatGPT to enhance students' writing skills, engagement, and feedback literacy", in *Edelweiss Applied Science and Technology*, 8, 4, pp. 2306-2319: <https://doi.org/10.55214/25768484.V8i4.1600>.
- Olson M. W., Raffeld P. (1987), "The effects of written comments on the quality of student compositions and the learning of content", in *Reading Psychology*, 8,4, pp. 273-293: <https://doi.org/10.1080/0270271870080404>.
- Ouyang F., Zheng L., Jiao P. (2022), "Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020", in *Education and Information Technologies*, 27, 1, pp. 7893-7925: <https://doi.org/10.1007/s10639-022-10925-9>.
- Page E. B. (2003), "Project essay grade: PEG", in Shermis M. D., Burstein J. C. (eds.), *Automated essay scoring: A cross-disciplinary perspective*, Routledge, New York, pp. 39-50.
- Page M. J., McKenzie J. E., Bossuyt P. M., Boutron I., Hoffmann T. C., Mulrow C. D., Shamseer L., Tetzlaff J. M., Akl E. A., Brennan S. E., Chou R., Glanville J., Grimshaw J. M., Hróbjartsson A., Lalu M. M., Li T., Loder E. W., Mayo-Wilson E., McDonald S., McGuinness L. A., Stewart L. A., Thomas J., Tricco C. A., Welch V. A., Whiting P., Moher D. (2021), "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews", in *BMJ*, 372, 71: <https://doi.org/10.1136/bmj.n71>.
- Patthey-Chavez G. G., Matsumura L. C., Valdés R. (2004), "Investigating the process approach to writing instruction in urban middle schools", in *Journal of Adolescent & Adult Literacy*, 47,6, pp. 462-476.
- Pérez A., McClain S. K., Roa A. F., Rosado-Mendinueta N., Trigos-Carrillo L., Robles H., Campo O. (2025), "Artificial intelligence applications in college academic writing and composition: A systematic review", in *Ikala*, 30, 1, pp. 1-37: <https://doi.org/10.17533/udea.ikala.355878>.
- Piaget J. (1953), *The origin of intelligence in the child*, Routledge, London.
- Powell K. C., Kalina C. J. (2009), "Cognitive and social constructivism: Developing tools for an effective classroom", in *Education*, 130, 2, pp. 241-250.
- Puentedura R. (2014), *Learning, technology, and the SAMR model: Goals, processes, and practice*: <http://www.hippasus.com/rpweblog/archives/2014/06/29/LearningTechnologySAMRModel.pdf>.
- Ranalli J., Hegelheimer V. (2022), "Introduction to the special issue on automated writing evaluation", in *Language Learning and Technology*, 26, 2, pp. 1-4: <http://doi.org/10125/73473>.
- Roscoe R. D., Brandon R. D., Snow D. L., MacNamara D. S. (2014), "Game-based writing strategy practice with the Writing Pal", in Pytash K. E., Ferdig R. E. (eds.), *Exploring technology for writing and writing instruction*, Information Science Reference, USA, pp. 1-20.
- Schmidt R. (1990), "The role of consciousness in second language learning", in *Applied Linguistics*, 11, 2, pp. 129-158: <https://doi.org/10.1093/applin/11.2.129>.
- Schmidt T., Strasser T. (2022), "Artificial Intelligence in foreign language learning and teaching: A CALL for intelligent practice", in *Anglistik: International Journal of English Studies*, 33, 1, pp. 165-184.
- Selinker L. (1992), *Rediscovering interlanguage*, Longman, London.
- *Seo J. -Y. (2024), "Exploring the educational potential of ChatGPT: AI-assisted narrative writing for EFL college students", in *Language Teaching Research Quarterly*, 43, 1, pp. 1-21: <https://doi.org/10.32038/ltrq.2024.43.01>.
- Shi H., Aryadoust V. (2024), "A systematic review of AI-based automated written feedback research", in *ReCALL*, 36, 2, pp. 187-209:

<https://doi.org/10.1017/S0958344023000265>.

- Sommers N. (1980), Revision strategies of student writers and experienced adult writers, in *College Composition and Communication*, 31, 4, pp. 378-388:
<https://doi.org/10.2307/356588>.
- Stevenson M., Phakiti A. (2019), “Automated feedback and second language writing”, in Hyland K., Hyland F. (eds.), *Feedback in second language writing: Contexts and issues*, Cambridge University Press, Cambridge, pp. 125-142:
<https://doi.org/10.1017/9781108635547.009>.
- *Strobl C., Menke-Bazhutkina I., Abel N., Michel M. (2024), “Adopting ChatGPT as a writing buddy in the advanced L2 writing class”, in *Technology in Language Teaching and Learning*, 6, 1, pp. 1-19: <https://doi.org/10.29140/TtL.V6n1.1168>.
- Su Y., Lin Y., Lai C. (2023), “Collaborating with ChatGPT in argumentative writing classrooms”, in *Assessing Writing*, 57, pp. 1-13:
<https://doi.org/10.1016/j.asw.2023.100752>.
- Swain M. (1993), “The output hypothesis: just speaking and writing aren’t enough” in *The Canadian Modern Language Review*, 50, 1, pp. 158-164:
<https://doi.org/10.3138/cmlr.50.1.158>.
- Teng M. F. (2024), “A systematic review of ChatGPT for English as a foreign language writing: Opportunities, challenges, and recommendations”, in *International Journal of TESOL Studies*, 6, 3, pp. 36-57: <https://doi.org/10.58304/ijts.20240304>.
- *Tsai C. -Y., Lin Y. -T., Brown I. K. (2024), “Impacts of ChatGPT-assisted writing for EFL English majors: Feasibility and challenges”, in *Education and Information Technologies*, 21, 7, pp. 22427-22445: <https://doi.org/10.1007/S10639-024-12722-Y>.
- *Tseng Y. -C., Lin Y. -H. (2024), “Enhancing English as a foreign language (EFL) learners’ writing with ChatGPT: A university-level course design”, in *Electronic Journal Of E-Learning*, 22, 2, pp. 78-97: <https://doi.org/10.34190/Ejel.21.5.3329>.
- Vojak C., Kline S., Cope B., McCarthy S., Kalantzis M. (2011), “New spaces on old places: An analysis of writing assessment software”, in *Computers and Composition*, 28, 2, pp. 97-111: <https://doi.org/10.1016/j.compcom.2011.04.004>.
- Vygotsky L. S. (1978), *Mind in society: the development of higher psychological processes*, Harvard University Press, Cambridge (USA)-London.
- *Wang D. (2024), “Teacher- versus AI-generated (Poe application) corrective feedback and language learners’ writing anxiety, complexity, fluency, and accuracy”, in *International Review of Research in Open and Distributed Learning*, 25, 3, pp. 37-55:
<https://doi.org/10.19173/irrodl.v25i3.7646>.
- Wang Y. -J., Shang H. -F., Briody P. (2013), “Exploring the impact of using automated writing evaluation in English as a foreign language university students’ writing”, in *Computer Assisted Language Learning*, 26, 3, pp. 234-257:
<https://doi.org/10.1080/09588221.2012.655300>.
- Wang S., Wang F., Zhu Z., Wang J., Tran T., Du Z. (2024), “Artificial intelligence in education: A systematic literature review”, in *Expert Systems with Applications*, 252, Part A, pp. 124-167: <https://doi.org/10.1016/j.eswa.2024.124167>.
- Warschauer M., Tseng W., Yim S., Webster T., Jacob S., Du Q., Tate T. (2023), “The affordances and contradictions of AI-generated text for writers of English as a second or foreign language”, in *Journal of Second Language Writing*, 63, pp. 1-7:
<https://doi.org/10.1016/j.jslw.2023.101071>.
- *Werdiningsih I., Indrawati I., Rusdin D., Ivone F. M., Basthomi Y., Zulfahreza (2024a), “Revolutionizing EFL writing: Unveiling the strategic use of ChatGPT by Indonesian master’s students”, in *Cogent Education*, 11, 1, pp. 1-18:
<https://doi.org/10.1080/2331186X.2024.2399431>.

- *Werdiningsih I., Marzuki, Rusdin D. (2024b), “Balancing AI and authenticity: EFL students’ experiences with ChatGPT in academic writing”, in *Cogent Arts and Humanities*, 11, 1, pp. 1-15: <https://doi.org/10.1080/23311983.2024.2392388>.
- Yang J., Jin H., Tang R., Han X., Feng Q., Jiang H., Zhong S., Yin B., Hu X. (2024), “Harnessing the power of LLMs in practice: A survey on ChatGPT and beyond”, in *ACM Transactions on Knowledge Discovery from Data*, 18, 6, pp. 1-32: <https://doi.org/10.1145/3649506>.
- Young J. C., Shishido M. (2023), “Investigating OpenAI’s ChatGPT potentials in generating chatbot’s dialogue for English as a Foreign Language Learning”, in *International Journal of Advanced Computer Science and Applications*, 14, 6, pp. 65-72: <http://dx.doi.org/10.14569/IJACSA.2023.0140607>.
- Zawacki-Richter O., Marín V. I., Bond M., Gouverneur F. (2019), “Systematic review of research on artificial intelligence applications in higher education – where are the educators?”, in *International Journal of Education Technology in Higher Education*, 16, 39, pp. 1-27: <https://doi.org/10.1186/s41239-019-0171-0>.
- Zimmermann B. J. (2000), “Attaining self-regulation: A social-cognitive perspective”, in Boekarts M., Pintrich P. R., Zeidner M. (eds.), *Handbook of Self-Regulation*, Academic Press, Orlando (FL), pp. 13-39.
- Zhu M., Wang C. (2025), “A systematic review of research on AI in language education: Current status and future implications”, in *Language Learning and Technology*, 29, 1, pp. 1-29: <https://hdl.handle.net/10125/73606>.

