NEXUS : ARTICULER PRATIQUES D'ENSEIGNEMENT ET RECHERCHES
NEXUS: CONNECTING TEACHING PRACTICE AND RESEARCH

L'usage de l'IAG pour l'enseignement/ apprentissage de la production écrite en FLE/FLS : une revue de portée

LÊ Hải Yến
Université de Langues et d'Études internationales
Université nationale du Vietnam à Hanoï

Résumé

Cette revue de portée examine les usages de l'intelligence artificielle générative (IAG) – notamment ChatGPT et outils apparentés – dans l'enseignement/apprentissage de la production écrite en FLE/FLS. Conformément à PRISMA-ScR, nous avons mené une recherche « toute source/tout type », puis appliqué des critères d'inclusion fondés sur le schéma PCC (Population-Concept-Contexte) avec exigence d'accès au texte intégral. Les travaux récents convergent vers plusieurs usages : révision de textes avec rétroaction générée par l'IAG, visant l'amélioration des dimensions linguistiques (lexique, syntaxe) et de la cohérence ; typologies de prompts et de réponses utiles à l'étayage de l'écriture (détection/explication d'erreurs, suggestions de correction) ; dispositifs de formation des enseignants/étudiants-enseignants mobilisant des textes produits par l'IAG pour concevoir l'enseignement de l'écrit (ingénierie de tâches, critères) ; études de cas documentant autonomie, motivation et appropriation raisonnée en FLE. Les bénéfices rapportés portent sur la correction linguistique, la rapidité de la rétroaction et le soutien à la révision. Les limites récurrentes concernent la dépendance aux prompts, la variabilité de la qualité des réponses et les enjeux éthiques (intégrité, déclaration/traçabilité de l'usage). Des lacunes persistent : rareté d'études comparatives rigoureuses par niveaux CECR, peu de protocoles longitudinaux et besoin d'alignement didactique-évaluatif. En conséquence, nous formulons des pistes de scénarisation (ingénierie de prompts, consignes métacognitives, transparence d'usage) et un agenda de recherche en vue d'une intégration responsable de l'IAG dans l'enseignement de la production écrite en FLE/FLS.

Mots-clés : FLE/FLS, production écrite, intelligence artificielle générative, ChatGPT, enseignement du français

Introduction

Au cours des dernières années, l'essor des modèles de langage de grande taille (LLM) et, plus largement, de l'intelligence artificielle générative (IAG) a profondément reconfiguré le paysage de l'enseignement-apprentissage des langues. Dans l'espace francophone, cette évolution pose avec acuité la question de la production écrite en français langue étrangère/seconde (FLE/FLS) : comment articuler les potentialités des outils génératifs (p. ex. ChatGPT, Claude, Gemini, DeepL Write, QuillBot) avec les exigences didactiques de l'écrit – au sens d'une compétence complexe qui mobilise orthographe et grammaire, lexique, cohérence textuelle, adéquation au genre, planification/révision et conscience métacognitive ?

Sur le plan didactique, l'IAG est susceptible d'intervenir à plusieurs niveaux de l'activité scripturale envisagée comme un processus d'écriture articulant planification, mise en texte et révision (Hayes, 2012) : idéation et planification (trouver des idées, organiser un plan, reformuler une consigne), rétroaction et révision (détection d'erreurs, explication, proposition d'améliorations), édition (paraphrase, simplification/complexification contrôlée, harmonisation stylistique) et évaluation formative (grilles d'autoévaluation accompagnées, pistes de réécriture). Ces usages de l'IAG, lorsqu'ils sont scénarisés et étayés, pourraient accélérer le cycle d'écriture et favoriser l'autonomie du/de la personne apprenante. À l'inverse, en l'absence de balises, ils exposent à divers risques : dépendance cognitive, affaiblissement des processus de planification et de révision, opacité des décisions produites par les systèmes et enjeux d'intégrité (traçabilité des aides, déclaration d'usage, plagiat non intentionnel).

Au-delà du seul contexte FLE/FLS, l'essor récent de l'IAG s'inscrit dans un mouvement plus large de transformation de l'enseignement des langues. Des travaux de cadrage montrent que la diffusion de ces outils reconfigure non seulement les pratiques pédagogiques mais aussi les questions d'évaluation, de recherche en acquisition des langues et de formation des enseignant·es (Chapelle, 2025). Dans le champ du CALL/ALAO/ELAO, l'IAG peut ainsi être envisagée comme un environnement de médiation offrant de nouvelles possibilités d'interaction, de rétroaction et d'apprentissage autonome, tout en soulevant des limites liées à son fonctionnement statistique, à ses biais et à ses conditions d'usage (Godwin-Jones, 2024). Ce double mouvement – ouverture d'affordances nouvelles et nécessité d'un encadrement critique – constitue un point d'ancrage essentiel pour lire les usages observés en production écrite.

Ce cadrage est d'autant plus important que plusieurs usages recensés dans le corpus relèvent de la révision textuelle et de la rétroaction corrective. Or, ces pratiques ne surgissent pas dans un vide théorique : elles prolongent des débats déjà bien établis sur le *written corrective feedback* en acquisition des langues et en didactique de l'écriture, débats relancés notamment à la suite des positions de Truscott et des synthèses critiques qui ont suivi (Ferris, 2012). Dans cette perspective, l'IAG ne doit pas être pensée uniquement comme un outil de correction ou de reformulation, mais comme

un médiateur susceptible d'agir sur différentes étapes du processus rédactionnel. Des travaux récents sur l'écriture en L2 invitent d'ailleurs à penser simultanément ses affordances et ses contradictions, en insistant sur la nécessité d'un accompagnement pédagogique centré sur la compréhension, l'accès, le prompting, la corroboration et l'intégration raisonnée des sorties générées par l'outil (Warschauer *et al.*, 2023).

Dans ce sous-champ plus étroit du FLE/FLS, les recherches portant sur l'usage de l'IAG dans l'enseignement/apprentissage de la production écrite restent récentes et hétérogènes : les dispositifs varient selon le niveau CECR (d'A1 à C1), les contextes (université, lycée, autoformation), la nature des genres discursifs (messages courts, écrits académiques, genres professionnels) et la place réservée à l'IAG (assistance ponctuelle ou intégrée dans une séquence didactique). Les indicateurs mobilisés ne sont pas stabilisés : certains travaux retiennent des scores analytiques (lexique, morphosyntaxe, cohérence), d'autres des mesures de complexité, de longueur, ou encore des marqueurs motivationnels et d'autonomie. Dans ce contexte mouvant, une revue de portée (*scoping review*) s'impose : il s'agit d'établir la carte des usages et des résultats rapportés, d'identifier les pratiques d'étayage et les garde-fous éthiques, et de faire émerger les lacunes qui structureront un agenda de recherche et d'innovation didactique en FLE/FLS.

Sur le plan didactique, la revue vise quatre résultats concrets pour la communauté FLE/FLS :

- une cartographie des usages de l'IAG dans l'écriture (rétroaction/édition, génération guidée, remédiation orthographique/grammaticale, évaluation formative), par niveaux CECR et contextes ;
- un répertoire d'étayages (ingénierie de prompts, consignes métacognitives, modalités de déclaration d'usage et de traçabilité) aligné sur les genres et critères d'évaluation de l'écrit en FLE ;
- un agenda de recherche précisant les besoins en devis (quasi-expérimentaux/longitudinaux), en mesures standardisées de la qualité textuelle et en politiques d'intégrité pour des environnements d'apprentissage outillés par l'IA ;
- une mise en évidence systématique des lacunes (*gaps*) dans l'état des preuves afin d'orienter les travaux futurs, conformément aux objectifs des revues de portée visant à cartographier les preuves et à identifier les lacunes (Arksey et O'Malley, 2005).

Les questions de recherche qui guident la revue sont les suivantes :

- QR1. Comment l'IAG est-elle mobilisée pour enseigner, apprendre et/ou évaluer la production écrite en FLE/FLS ?
- QR2. Quels résultats d'apprentissage sont rapportés (exactitude, cohérence, complexité, longueur, motivation, autonomie) et comment sont-ils mesurés ?
- QR3. Quelles formes d'étayage et de gouvernance éthique (transparence d'usage, intégrité académique, usages responsables) sont décrites ?
- QR4. Quelles lacunes prioritaires la cartographie met-elle en évidence (populations/CECR sous-étudiées, genres/tâches non couverts, mesures/devis

manquants, risques éthiques peu documentés) et quelles priorités de recherche en découlent ?

Enfin, cette revue se veut utile tant aux enseignant·es qu'aux décideurs et concepteurs de dispositifs en FLE/FLS. Elle propose des repères opérationnels pour intégrer l'IAG sans dévoyer les apprentissages ciblés par le CECR : faire de l'IA un instrument d'étayage de la planification et de la révision plutôt qu'un substitut à la pensée écrite ; renforcer les pratiques d'évaluation (barèmes, critères, preuves de révision) ; instaurer des règles d'usage transparentes au bénéfice de l'intégrité et du développement de compétences durables. Les sections suivantes détaillent la méthode avant de présenter la cartographie des études, la synthèse thématique et les implications pédagogiques et scientifiques.

1. Méthodologie

La présente contribution prend la forme d'une revue de portée conforme aux recommandations PRISMA-ScR (Tricco *et al.*, 2018), dont l'objectif est de cartographier l'étendue et la nature des preuves disponibles sur les usages de l'IAG dans l'enseignement/apprentissage de la production écrite en FLE/FLS, ainsi que d'en identifier les principales lacunes, plutôt que d'estimer des effets (Levac *et al.*, 2010). L'éligibilité des documents est guidée par le cadre PCC (Population–Concept–Contexte). Conformément au mandat d'une *scoping review*, aucune évaluation critique JBI de la qualité méthodologique n'est conduite ; il s'agit ici de proposer une vue d'ensemble structurée, transparente et reproductible des dispositifs, des usages et des résultats rapportés.

1.1. Critères d'éligibilité (PCC) et double corpus

Nous distinguons deux ensembles complémentaires. Le corpus principal regroupe les études empiriques qui impliquent des apprenant·e·s de FLE/FLS réalisant des tâches de production écrite, avec des données observables (textes, scores, journaux d'interaction, entretiens, etc.). Le corpus complémentaire rassemble les retours d'expérience centrés sur les enseignant·e·s et les articles de cadrage/point de vue qui documentent la scénarisation, l'étayage et les enjeux éthiques, mais sans échantillon d'apprenants. Le Concept exige un lien explicite avec l'écrit (idéation, rétroaction, révision, co-écriture, évaluation formative, remédiation orthographique). Le Contexte est limité au FLE/FLS, tous niveaux CECR et environnements (université, secondaire, formation continue, migration, autoformation).

1.2. Sources d'information et stratégie de recherche

Pour tenir compte des contraintes d'accès, nous adoptons une stratégie OA-first et interrogeons exclusivement des portails et archives ouverts : Google Scholar, HAL, OpenEdition (dont ALSIC), DOAJ, ERIC (accès libre), Cairn pour les contenus ouverts, des revues thématiques en *open access* (p. ex. *RITPU, PARÉO*), ainsi que theses.fr/TEL pour les thèses et mémoires. Une recherche boule de neige (*backward/forward*) complète le dispositif à partir des bibliographies des études retenues. Les chaînes sont bilingues FR/EN afin de couvrir FLE/FLS, l'écrit et l'IAG ; à titre d'exemples :

- FR : ("français langue étrangère" OR FLE OR FLS) AND ("production écrite" OR écriture) AND ("IAG" OR "intelligence artificielle générative" OR ChatGPT) ;
- EN : ("French as a foreign/second language" OR FLE OR FLS) AND (writing OR "written production") AND ("generative AI" OR "large language model" OR ChatGPT).

Aucune restriction d'année n'est appliquée en amont ; le filtre « texte intégral disponible » intervient au criblage. Note de transparence : les bases sur abonnement (Scopus, Web of Science, etc.) ne sont pas utilisées pour l'accès aux textes ; lorsqu'elles servent au repérage, seules les versions OA correspondantes (HAL, sites éditeurs en libre accès, TEL, etc.) sont incluses.

1.3. Processus de sélection

Le criblage s'est déroulé en deux étapes. D'abord, les titres et résumés ont été examinés au regard du PCC et de la disponibilité du texte intégral. Ensuite, les textes complets retenus ont été évalués pour décider de l'inclusion et du classement en corpus principal ou corpus complémentaire. Le résumé du flux de sélection rend compte des effectifs à chaque étape (identification, dédoublonnage, exclusions successives, inclusions finales) ainsi que des motifs d'exclusion.

1.4. Extraction des données (*data-charting*)

Les données ont été extraites dans une grille normalisée. Pour le corpus principal, la grille comporte 13 champs : Citation (APA7), Année, Pays/Contexte, Population/Niveau CECR, Type d'apprenants, Outil IAG, Tâches d'écriture, Conception de l'étude, Durée/Contexte pédagogique, Indicateurs/Mesures, Principaux résultats, Enjeux éthiques/Limites, Notes. Pour le corpus complémentaire, une grille réduite saisit la référence, le contexte, la focalisation (enseignant/cadrage), les outils IAG, les axes (étayage/éthique), les messages clés et les notes. Par principe de preuve, toute information reportée est adossée à un extrait (≤ 25 mots) avec page ; lorsqu'un élément

n'est pas fourni par la source, la mention « non précisé » est conservée, assortie d'un équivalent explicitement indiqué si pertinent (p. ex. « niveau intermédiaire faible »).

1.5. Méthode de synthèse

La synthèse procède en deux temps pour le corpus principal : une cartographie descriptive (pays, niveaux CECR, outils, tâches, designs, durées) puis une synthèse thématique des usages (idéation, rétroaction, révision, co-écriture, évaluation formative), des modalités d'étayage (guidage de prompts, consignes métacognitives, déclaration d'usage de l'IAG) et des indicateurs/résultats (précision, complexité, cohérence, motivation, autonomie, etc.). Le corpus complémentaire fait l'objet d'une synthèse narrative orientée vers la scénarisation pédagogique, la littératie IAG et les enjeux éthiques. Aucune méta-analyse n'est envisagée compte tenu de l'hétérogénéité attendue.

1.6. Intégrité et limites méthodologiques

Nous garantissons l'absence d'invention de données et la correspondance stricte entre les citations dans le texte et la bibliographie au format APA 7. Toutes les sources mobilisées ont été consultées en texte intégral et le journal de recherche ainsi que les grilles d'extraction sont archivés pour vérification. Les limites anticipées tiennent à l'hétérogénéité des dispositifs et des mesures dans un champ en évolution rapide, ainsi qu'à un biais linguistique possible (FR/EN). Ces limites seront prises en compte lors de l'interprétation des résultats.

2. Résultats

2.1. Sélection des études

Au total, 25 enregistrements ont été identifiés lors de la phase d'identification. Après élimination des doublons bibliographiques, 22 notices ont été conservées pour le criblage titres/résumés. À cette étape, 2 enregistrements ont été écartés (doublons de contenu), de sorte que 20 documents ont été soumis à la lecture du texte intégral.

La lecture intégrale a abouti à une inclusion totale de 20 documents, répartis en deux ensembles complémentaires conformément à notre protocole PCC : 13 études dans le corpus principal – impliquant des apprenant·e·s et des tâches de production écrite avec données empiriques – et 7 études classées dans le corpus complémentaire (retours d'expérience ou articles de cadrage centrés sur l'enseignant/la conception), conservées pour éclairer la scénarisation et l'éthique, mais non agrégées à la synthèse des effets sur l'écrit.

Le flux PRISMA se présente ainsi :

- Identifiés : 25 ; Doublons retirés : 3 → Restants : 22

- Criblage titres/résumés : 22 → Exclus : 2 (doublons de contenu)
- Textes intégraux évalués : 20 → Inclus : 20 (= 13 dans le corpus principal + 7 dans le corpus complémentaire)
- Exclusions au texte intégral (motifs PCC) : 0 hors-concept ; 0 hors-contexte ; 0 indisponible.
  (Remarque : les 7 études du corpus complémentaire ne sont pas exclues de la revue ; elles sont reclassées pour un usage non agrégé dans la Discussion.)

La recherche n'imposait pas de borne temporelle ; les études retenues sont néanmoins publiées depuis 2023, en phase avec l'essor récent des IAG. Les textes sont en français, anglais et vietnamien, sans exclusion de thèses/mémoires ou preprints à l'étape du texte intégral.

2.2. Caractéristiques des études (corpus principal)

2.2.1. Vue d'ensemble

Le corpus principal comprend 13 études publiées très récemment (2024 : 6 ; 2025 : 7). Les contextes sont variés, couvrant l'Afrique du Nord (p. ex. Algérie : Bechiri, 2024 ; Heddouche, 2024 ; Khelafi et Bourkaib Saci, 2024 ; Maroc : El Wahabi et Belhaj, 2024), l'Europe (France, contexte migration : Schlemminger, 2025a), l'Asie (Turquie : Dündar et Aydoğdu, 2024 ; Thaïlande : Marchal et Phoungsub, 2025 ; Vietnam : Do, 2025a), l'Amérique du Nord (Canada : Holmes, 2024) et l'Amérique latine (Colombie : Molina Zapata et Montoya Estrada, 2025). Les publics vont des adultes débutants A1 (migration) aux étudiants de licence/master (avec plusieurs niveaux CECR rapportés), et incluent un échantillon large multi-niveaux (A1 à C1) dans l'enquête thaïlandaise (Marchal et Phoungsub, 2025).

2.2.2. Outils d'IAG et types de tâches d'écriture

L'outil dominant est ChatGPT (12/13 études). On recense aussi Perplexity (Bechiri, 2024) et des agents conversationnels conçus spécifiquement pour l'expression écrite (Dündar et Aydoğdu, 2024). Les tâches portent majoritairement sur la révision/autorévision (p. ex. Heddouche, 2024 ; Do, 2025a), la rétroaction corrective (détection/explication d'erreurs), l'argumentation/dissertation (El Wahabi et Belhaj, 2024), le résumé (Heddouche, 2024) et des écritures guidées contextualisées (migration : Schlemminger, 2025a ; Alliance française : Molina Zapata et Montoya Estrada, 2025).

2.2.3. Conceptions et durées d'intervention

Les conceptions d'étude sont variées : les méthodes mixtes sont les plus fréquentes (5/13) ; des démarches de recherche-action sont également rapportées (3/13). On

relève des dispositifs quasi-expérimentaux avec groupe témoin (p. ex. Bechiri, 2024 ; Molina Zapata et Montoya Estrada, 2025 : évaluations avant/après). La durée est explicitement chiffrée dans 8/13 études (p. ex. séquence de 4 mois avec écrits hebdomadaires et intensification en mai : El Wahabi et Belhaj, 2024 ; 3 séances intensives × 3 h : Bechiri, 2024), tandis que 5/13 ne précisent pas le calendrier (travaux de corpus ou interventions à pas fixe non rapportés).

2.2.4. Indicateurs et mesures d'effets

Les études du corpus mobilisent des indicateurs variés pour documenter les effets de l'IAG sur l'apprentissage de l'écrit en FLE/FLS. Les plus fréquemment relevés concernent la précision linguistique, notamment l'orthographe, la grammaire et la détection des erreurs (Bechiri, 2024 ; Do, 2025a ; Li, 2025). D'autres travaux prennent également en compte la complexité textuelle ou argumentative (El Wahabi et Belhaj, 2024 ; Khelafi et Bourkaib Saci, 2024), la cohérence ou l'organisation du texte (Bechiri, 2024 ; El Wahabi et Belhaj, 2024), ainsi que certaines dimensions psychopédagogiques telles que la motivation ou l'autonomie (Heddouche, 2024 ; Elmouhtadi *et al.*, 2025 ; Marchal et Phoungsub, 2025).

Les modalités de mesure demeurent hétérogènes. Certaines études s'appuient sur des comparaisons pré/post ou sur des tests d'hypothèse (Bechiri, 2024 ; Molina Zapata et Montoya Estrada, 2025), tandis que d'autres recourent à des approches plus qualitatives ou descriptives, fondées sur des grilles d'évaluation, des codages d'erreurs, des questionnaires, des journaux de bord, des entretiens ou des productions d'apprenants (Do, 2025a ; Heddouche, 2024 ; Elmouhtadi *et al.*, 2025 ; Marchal et Phoungsub, 2025).

2.2.5. Encadrement pédagogique (aperçu descriptif)

La majorité des études adossent l'IA à un étayage : guidage de prompts (contexte, rôle, registre, longueur, exemples), consignes métacognitives (justifier les choix, comparer les versions initiales/révisées) et déclaration d'usage. Les dispositifs en contexte réel (migration, Alliance française, universités) privilégient une intégration curriculaire avec alternance présentiel/numérique, parfois couplée à des outils d'analyse (p. ex. FLElex pour le niveau lexical en migration).

2.3. Synthèse thématique (corpus principal)

2.3.1. Usages de l'IAG pour l'écrit

Dans le corpus principal, l'IAG est mobilisée comme levier de processus d'écriture, et non comme substitut d'auteur. Les usages dominants relèvent de la rétroaction

et de la révision/autorévision : les apprenant·e·s produisent une version initiale, sollicitent l'IA sous consignes, puis comparent les versions selon des critères explicites (Bechiri, 2024 ; Heddouche, 2024 ; Do, 2025a). D'autres scénarios ciblent l'idéation/planification – génération d'idées, de plans ou de connecteurs – en particulier pour des écrits argumentatifs (El Wahabi et Belhaj, 2024). Selon le public et le contexte, l'IA sert aussi de support de remédiation orthographique/grammaticale et de résumé/paraphrase, utile pour des profils débutants ou auprès de publics primo-arrivants (Schlemminger, 2025a). Enfin, quelques dispositifs l'adossent à l'évaluation formative, où des grilles et critères sont posés en amont, l'IA fournissant des pistes que l'enseignant·e et/ou l'apprenant·e valident (Dündar et Aydoğdu, 2024 ; Bechiri, 2024).

2.3.2. Résultats rapportés et mesures

La cartographie des indicateurs montre un accent net sur l'exactitude linguistique, qui constitue la dimension la plus fréquemment documentée dans le corpus : les études suivent surtout les erreurs, l'orthographe et la grammaire au moyen de grilles ou de codages d'erreurs (Bechiri, 2024 ; Do, 2025a, entre autres). Les dimensions de complexité/argumentation et de cohérence/organisation sont moins fréquemment documentées, mais, lorsqu'elles sont présentes, on observe un meilleur usage des connecteurs, une structuration plus nette (introduction, développement, conclusion) et une articulation plus solide des arguments (El Wahabi et Belhaj, 2024 ; Khelafi et Bourkaib Saci, 2024). Les marqueurs de motivation/autonomie demeurent plus ponctuellement rapportés, essentiellement via des données qualitatives (journaux, entretiens, autoévaluations) (Heddouche, 2024 ; Schlemminger, 2025a).

Côté méthodes, les dispositifs pré/post et les tests statistiques restent minoritaires, la majorité des travaux procédant par comparaison descriptive avant/après appuyée sur des grilles (DELF adaptées ou internes) et des exemples textuels. Globalement, les effets positifs documentés concernent surtout la réduction d'erreurs et une meilleure organisation dès lors que l'étayage est rigoureux ; en parallèle, plusieurs études signalent la variabilité du feedback de l'IA (dont des cas de surcorrection), d'où la nécessité d'un tri critique par l'apprenant·e et/ou l'enseignant·e (Do, 2025a ; Bechiri, 2024).

2.3.3. Étayage pédagogique et gouvernance éthique

Les études convergent vers l'idée que la valeur didactique de l'IA dépend du degré d'étayage. Les dispositifs efficaces combinent guidage des prompts (contexte, rôle, registre, longueur attendue, exemples) et consignes métacognitives (justifier des choix, expliciter une stratégie, comparer versions) (Schlemminger, 2025a ; Dündar et Aydoğdu, 2024). L'intégration curriculaire se fait dans des cadres authentiques – université, lycée, Alliance française, formation de migrants – souvent en présence/distanciel alternés, parfois avec des outils d'analyse pour objectiver le niveau lexical ou

la structure (Bechiri, 2024 ; Molina Zapata et Montoya Estrada, 2025 ; Schlemminger, 2025a). Sur le plan éthique, les études mettent en avant des principes de transparence d'usage, l'idée que l'IA ne doit pas se substituer à la compétence scripturale, ainsi que la nécessité d'une formation aux compétences numériques pour limiter la dépendance et les imprécisions ou la surcorrection (Heddouche, 2024 ; Schlemminger, 2025a, 2025b). Dans ces conditions, l'IA est positionnée comme tuteur/guide susceptible d'accroître l'autonomie sans fragiliser l'intégrité académique.

2.3.4. Lacunes cartographiées et priorités de recherche

Quatre lacunes saillantes émergent de la cartographie.

- Une faible standardisation des mesures de cohérence et de complexité, alors qu'elles sont centrales pour l'écrit académique ; un besoin apparaît de rubriques/critères plus partagés et de procédures d'annotation stabilisées (El Wahabi et Belhaj, 2024 ; Khelafi et Bourkaib Saci, 2024).
- Des devis pré/post avec groupe témoin ou longitudinaux encore rares, limitant l'estimation de la durabilité des effets (Bechiri, 2024 ; Molina Zapata et Montoya Estrada, 2025).
- Des rapports encore épars portent sur les dimensions psycho-stratégiques, notamment la motivation et l'autonomie, qui restent documentées dans un nombre limité d'études et appellent des instruments plus stables (Heddouche, 2024 ; Elmouhtadi *et al.*, 2025).
- Une hétérogénéité persistante du reporting d'étayage et d'éthique (niveau de guidage des prompts, modalités de transparence d'usage, critères en présence de l'IA), qui complique les comparaisons ; des lignes de rapport par type de tâche et niveau CECR, assorties d'un degré de contrôle explicite, seraient utiles pour cumuler les connaissances.

2.4. Études complémentaires – apports à la scénarisation et à l'éthique

Les sept études du corpus complémentaire n'impliquent pas des dispositifs expérimentaux avec apprenant·e·s ; elles prennent la forme de cadres conceptuels, guides de pratiques de classe, retours d'expérience ou réflexions éthiques centrées sur l'enseignant·e et la conception. Conformément à PRISMA-ScR, elles sont retenues pour éclairer l'angle didactique et les enjeux d'usage, mais non agrégées aux effets d'apprentissage rapportés par le corpus principal.

2.4.1. Panorama thématique

À l'échelle du champ, ces contributions convergent sur quatre axes utiles à la scénarisation de l'écrit en FLE/FLS :

- Ingénierie de prompts (rôle, registre, contraintes de longueur, exemples), avec des protocoles de guidage progressif (du prompt « naïf » vers le prompt « expert ») ;
- Orchestration en classe (présentiel/hybride), y compris la position de l'IA comme tuteur et l'articulation avec pairs/enseignant ;
- Critères et outils d'évaluation formative (grilles critériées adaptées aux genres, explicitation de la cohérence/organisation) ;
- Gouvernance éthique : déclaration d'usage de l'IAG, intégrité académique (limiter la sur-/sous-correction, traçabilité des sources), équité d'accès et littératie numérique.

2.4.2. Pistes de scénarisation transférables

Les textes du corpus complémentaire proposent des séquences-types pour l'idéation, la rétroaction et la révision : démarrer par des consignes de prompt contrôlées (contexte, objectif, genre), exiger une justification métacognitive des modifications et alterner interaction IA ↔ pairs/enseignant. Plusieurs sources recommandent de dissocier les rôles (p. ex. un prompt « détecter-expliquer » pour l'analyse d'erreurs, puis un prompt « réécrire-selon-critères » pour la version révisée), ce qui facilite l'apprentissage des critères (cohérence, progression thématique, connecteurs) et réduit la dépendance.

2.4.3. Gouvernance éthique et littératie IAG

Ces études formulent un socle commun : transparence d'usage (déclarer quand et comment l'IA est utilisée), non-substitution d'auteur et formation à la vérification (sources, faits, droits). Elles signalent des risques récurrents (qualité inconstante, sur-correction, dérive de dépendance, inégalités d'accès) et suggèrent des garanties procédurales : consignes explicites, journal de traces, double regard (IA + pairs/enseignant) et critères publics. L'ensemble cadre avec les exigences de rédaction d'une revue de portée (PRISMA-ScR), qui vise à mettre en évidence ce qui est décrit dans la littérature et ce qui manque encore.

2.4.4. Articulation avec le corpus principal

Pris comme réservoir de scénarios et de garde-fous, le corpus complémentaire renforce trois messages du corpus principal :

- l'efficacité de l'IA dépend d'un étayage explicite (prompts + métacognition) ;
- les critères (cohérence/organisation) gagnent à être exposés et entraînés en amont de la révision ;
- la gouvernance (déclaration d'usage, intégrité) doit être posée avant l'usage en évaluation.

Ces apports n'ajoutent pas de preuves d'effet, mais outillent l'implémentation des dispositifs du corpus principal et préparent la discussion des lacunes (2.3.4) et des implications pédagogiques.

3. Discussion

Dans l'ensemble, les études retenues convergent vers l'idée que l'IAG (principalement ChatGPT) peut soutenir l'amélioration de la production écrite en FLE/FLS, à condition d'un étayage didactique explicite (étayage, consignes, contrôle de l'usage) et d'une conscience métalinguistique suffisante chez les apprenant.e.s. Plusieurs résultats empiriques mettent en évidence des gains sur l'exactitude linguistique, la planification du texte et la révision – mais ces effets demeurent conditionnels à la qualité des requêtes (prompting) et au rôle médiateur de l'enseignant (Holmes, 2024 ; Mehrabi et Haghgou, 2025 ; Marchal et Phoungsub, 2025 ; Dündar et Aydoğdu, 2024).

3.1. Efficacité et conditions d'efficacité

Des améliorations significatives entre pré-test et post-test sont rapportées lorsque ChatGPT est intégré à des tâches d'écriture contrôlées, avec consignes et critères transparents (Mehrabi et Haghgou, 2025). Les données descriptives de classes montrent par ailleurs un usage fréquent de l'IA pour la correction, l'enrichissement lexical et la révision (Marchal et Phoungsub, 2025). Toutefois, la qualité des réponses de ChatGPT est variable et peut chuter lorsque les requêtes sont imprécises – ce qui renvoie à la nécessité d'apprendre à formuler des prompts et à vérifier les propositions de l'outil (Holmes, 2024). Dans des dispositifs arrimés aux descripteurs du CECR (par ex. critères DELF pour la structuration du plan), on observe des effets positifs sur l'organisation textuelle et l'appropriation de critères d'évaluation (Dündar et Aydoğdu, 2024). Ces constats rejoignent les analyses plus théoriques ou programmatiques qui plaident pour un usage raisonné de l'IAG, orienté compétences, évitant la délégation totale et intégrant un guidage explicite (Schlemminger, 2025b ; Nieto Escobar et Nadim, 2024).

3.2. Rôle de l'enseignant et de l'étayage

Plusieurs études de terrain soulignent qu'une utilisation autonome « pure » expose à des limites (feedback partiellement adéquat, simplifications excessives, tensions avec la complexité textuelle), tandis qu'un suivi pédagogique (consignes métacognitives, vérification humaine, critères partagés) favorise des gains plus robustes et durables (Li, 2025 ; Lorrillard, 2025 ; Rinck, 2025). Dans les cours où l'enseignant structure les étapes (idéation → rédaction → révision outillée → justification des corrections), les

apprenant·es développent mieux la vigilance métalinguistique et l'autonomie régulée plutôt qu'une dépendance à l'outil (El Wahabi et Belhaj, 2024 ; Bechiri, 2024). En évaluation formative, la combinaison prompt + correction humaine est recommandée pour sécuriser la qualité du retour (Do, 2025b), rejoignant l'appel à l'expertise enseignante pour estimer la pertinence des corrections en contexte professionnel (Constantin, 2023).

3.3. Effets hétérogènes selon tâches visées

Les effets rapportés de l'IAG apparaissent variés selon les tâches proposées et les compétences ciblées. Dans plusieurs études, les bénéfices sont surtout documentés sur l'exactitude linguistique, la révision et l'organisation du texte, notamment lorsque l'outil est mobilisé pour détecter des erreurs, suggérer des corrections, proposer des reformulations ou aider à structurer un plan (Holmes, 2024 ; Mehrabi et Haghgou, 2025 ; Dündar et Aydoğdu, 2024). D'autres travaux mettent davantage en évidence des effets sur l'argumentation, la créativité, la motivation ou l'autonomie, lorsque l'IA est intégrée à des tâches plus ouvertes ou à des dispositifs d'écriture guidée et réflexive (El Wahabi et Belhaj, 2024 ; Marchal et Phoungsub, 2025 ; Schlemminger, 2025a). Dans cette perspective, l'IA n'agit pas de manière univoque sur l'apprentissage de l'écrit : ses apports dépendent étroitement de la nature de la tâche, des objectifs visés et du degré d'étayage pédagogique.

L'étude de Molina Zapata et Montoya Estrada (2025) constitue, à cet égard, un cas particulièrement intéressant. Alors que plusieurs travaux du corpus font état de bénéfices plus nets dans les groupes ou dispositifs avec IA, cette recherche met en évidence une répartition plus nuancée des progrès selon les compétences observées : le groupe contrôle obtient de meilleurs résultats en compréhension orale et en production écrite, tandis que le groupe expérimental progresse davantage en compréhension écrite et en production orale. Un tel résultat invite à éviter toute lecture linéaire des effets de l'IA et à considérer plus finement l'articulation entre tâche, compétence visée et scénario pédagogique. Il renforce ainsi l'idée que l'efficacité didactique de l'IA dépend moins de sa seule présence que de son alignement avec les objectifs du cours et avec les activités réellement proposées aux apprenant·es.

3.4. Dimension éthique et littératie numérique

La littérature insiste sur la transparence d'usage, la déclaration des recours à l'IA et le développement d'une littératie éthique (plagiat, citations, propriété intellectuelle, biais) intégrée aux scénarios d'écriture (Schlemminger, 2025b ; Lorrillard, 2025). Les enquêtes auprès d'enseignant·es mettent en avant des bénéfices perçus (gain de temps, idées, correction d'erreurs non détectées) mais aussi la nécessité d'établir des règles

de cours et d'articuler l'IA avec des tâches authentiques et évaluations appropriées (Do, 2025b ; Rinck, 2025).

3.5. Convergences et divergences

- Convergences : potentiel de l'IAG pour soutenir la qualité linguistique et la planification ; exigence d'un étayage et d'une vérification humaine ; intérêt d'outils/critères CECR pour cadrer l'usage.
- Divergences : amplitude des gains selon la tâche (argumentation vs révision locale), le niveau CECR et la durée ; qualité de l'output IA fluctuante en fonction de la requête ; transférabilité limitée lorsque les dispositifs ne sont pas alignés au curriculum (Holmes, 2024 ; Mehrabi et Haghgou, 2025 ; Dündar et Aydoğdu, 2024 ; Molina Zapata et Montoya Estrada, 2025 ; Lorrillard, 2025).

3.6. Limites des preuves disponibles

Plusieurs études empiriques signalent des tailles d'échantillon réduites, des durées brèves, des plans quasi-expérimentaux sans assignation aléatoire, ou des mesures principalement centrées sur l'exactitude au détriment d'indicateurs de complexité et cohérence (Mehrabi et Haghgou, 2025 ; Do, 2025a ; Khelafi et Bourkaib Saci, 2024). Les travaux du corpus complémentaire (analyses, états des lieux, positions) sont utiles pour baliser l'éthique et la didactisation, mais n'apportent pas des preuves causales – d'où l'intérêt de lire leurs recommandations comme des conditions d'implémentation à tester dans des dispositifs contrôlés (Schlemminger, 2025b ; Nieto Escobar et Nadim, 2024 ; Constantin, 2023 ; Rinck, 2025).

3.7. Implications pédagogiques

Les résultats plaident pour des dispositifs où l'IA agit comme outil d'étayage régulé :

- pré-écriture (génération de plans, brainstorming guidé) ;
- rétroaction ciblée sur des micro-objectifs (orthographe, accords, connecteurs) avec contrôle enseignant ;
- révision justifiée par l'apprenant (métalangage, critères partagés) ;
- traçabilité de l'usage (journal de prompts/réponses) et déclaration dans les travaux évalués (Holmes, 2024 ; Dündar et Aydoğdu, 2024 ; Marchal et Phoungsub, 2025 ; Rinck, 2025).

L'alignement au CECR (niveau-cible explicite B1/B1+ et tâches calibrées), l'usage d'échelles analytiques (grilles critériées de production écrite – p. ex. critères DELF convertis en checklist d'autoévaluation) et une formation courte au « prompting » (atelier + modèles de prompts + accompagnement à l'ajustement) apparaissent comme

des leviers pour transformer des gains locaux en progrès global d'écriture (Dündar et Aydoğdu, 2024, p. 508-509 ; Mehrabi et Haghgou, 2025, p. 28-29).

3.8. Pistes de recherche

Les travaux suggèrent, d'une part, de renforcer la robustesse des devis par des études menées sur des échantillons plus larges et dans des contextes plus diversifiés, afin d'améliorer la généralisabilité des résultats (p. ex. échantillon limité à un seul établissement) (El Wahabi et Belhaj, 2024).

D'autre part, plusieurs contributions invitent à documenter les effets au-delà d'une temporalité courte, en examinant des effets à plus long terme (durée possiblement insuffisante pour observer des impacts durables) (El Wahabi et Belhaj, 2024).

Sur le plan des mesures, des travaux plaident pour une formalisation/standardisation de critères d'évaluation : en construisant une liste de critères précis pour évaluer l'outil et établir une typologie des problèmes rencontrés (Lorrillard, 2025) ; en s'appuyant sur des critères explicites servant à analyser des productions et à soutenir des activités de réécriture/comparaison (Rinck, 2025).

Concernant la différenciation par niveaux CECR et par types de tâches, certains textes montrent la possibilité de paramétrer les activités (exemples d'invites et de tâches) en précisant le niveau CECR et la nature du genre/objectif, ce qui ouvre la voie à des comparaisons plus systématiques entre tâches et niveaux (Nieto Escobar et Nadim, 2024). De plus, une étude centrée sur des apprenants intermédiaires souligne explicitement que les résultats ne permettent pas d'inférer ce qui se passerait chez des débutants ou des avancés, et appelle donc des recherches comparatives selon le niveau (El Wahabi et Belhaj, 2024).

Enfin, sur l'axe éthique/gouvernance, des auteurs insistent sur la nécessité d'une réglementation/encadrement en milieu éducatif pour éviter les usages frauduleux, tout en privilégiant une formation à un usage éthique et responsable plutôt qu'une interdiction (Nieto Escobar et Nadim, 2024).

En complément, des propositions portent sur le développement de protocoles d'utilisation « clairs et progressifs », incluant le calibrage de la fréquence/moment d'interaction et une progression des tâches (de la correction d'erreurs élémentaires vers la révision de textes plus complexes) (Li, 2025).

4. Limites de la revue

Cette revue cartographie le paysage des usages de l'IAG pour la production écrite en FLE/FLS sur la base d'un corpus sélectionné selon PRISMA-ScR, avec extraction normalisée (13 champs) et distinction entre corpus principal et corpus complémentaire. Nous soulignons ci-dessous les limites principales du processus de revue (et non

des études elles-mêmes), conformément à l'item 20 de la checklist PRISMA-ScR demandant d'expliciter les limites du processus de revue.

D'abord, la couverture des sources retenues – avec l'exigence d'un accès au texte intégral et l'exclusion des bases sous abonnement – expose la revue à un biais de disponibilité : des études potentiellement pertinentes mais inaccessibles au moment du criblage ont pu être écartées.

Ensuite, un biais linguistique et de champ demeure possible : le ciblage FLE/FLS et les langues de travail de la revue restreignent l'horizon de la cartographie et la transférabilité vers d'autres contextes.

De plus, l'hétérogénéité des devis et des mesures (quasi-expérimentations, études de cas, enquêtes ; indicateurs d'exactitude, de complexité, de cohérence, de motivation, etc.) empêche toute méta-analyse et limite les comparaisons quantitatives directes ; la synthèse appropriée reste donc descriptive/thématique.

En outre, la temporalité constitue une limite transversale : le champ des LLM/GenAI évolue rapidement (versions de modèles, politiques d'usage), si bien que des observations peuvent devenir périssables malgré une recherche récente.

Enfin, prises ensemble, ces limites impliquent que les résultats doivent être compris comme une cartographie fidèle mais non exhaustive de la littérature accessible sur la période considérée, sans inférence causale ni estimation d'effet.

Conclusion

Pris dans leur ensemble, les travaux retenus dressent une cartographie cohérente des usages de l'IAG pour la production écrite en FLE/FLS : le potentiel de l'outil se manifeste surtout sur l'exactitude linguistique et la planification/révision, à condition d'un étayage pédagogique explicite (prompts guidés, critères partagés alignés CECR, vérification humaine) et d'une conscience métalinguistique et éthique développée chez les apprenant·es. Toutefois, l'hétérogénéité des devis et des mesures, l'instabilité technologique des LLM et la disponibilité sélective des sources invitent à interpréter ces constats comme des tendances cartographiées plutôt que comme des estimations d'effet.

Sur le plan didactique, l'IA gagne à être intégrée comme outil d'étayage régulé : en amont (idéation, plan), pendant (rétroaction ciblée et traçable) et en aval (révision justifiée par l'étudiant·e), dans un dispositif aligné aux objectifs de cours et aux descripteurs du CECR. Sur le plan de la recherche, des essais contrôlés de plus longue durée, des mesures harmonisées (exactitude–complexité–cohérence, dimensions affectives), une traçabilité systématique des interactions avec l'IA et des études différenciées selon niveaux CECR et genres d'écriture constituent des priorités. En somme, cette revue met en évidence un levier didactique prometteur : l'IAG peut accélérer l'apprentissage de l'écrit lorsqu'elle est conçue, encadrée et évaluée comme un composant du scénario d'enseignement-apprentissage, et non comme un substitut à la compétence scripturale.

Bibliographie

Arksey, H. et O'Malley, L. (2005). Scoping studies: Towards a methodological framework. International Journal of Social Research Methodology, *8*(1), 19-32. https://doi.org/10.1080/1364557032000119616

Bechiri, C. (2024). Intégration de l'intelligence artificielle dans la classe de FLE : approches et pratiques pour l'amélioration de l'écrit à l'université de Skikda. *Ziglôbitha – Revue des Arts, Linguistique, Littérature & Civilisations*, *3*(10), 139-148.

Chapelle, C. A. (2025). Generative AI as game changer: Implications for language education. *System, 132*, 103672. https://doi.org/10.1016/j.system.2025.103672

Constantin, F. (2023). ChatGPT – Learning accelerator or demolisher of foreign language teaching and learning? An empirical study on Business French. *The Annals of the University of Oradea, Economic Sciences*, *32*(2), 225-238. https://doi.org/10.47535/1991auoes32(2)022

Do, T. B. T. (2025a). Feedback quality of ChatGPT in revising French texts. *HNUE Journal of Science: Educational Sciences*, *70*(2), 33-44. https://doi.org/10.18173/2354-1075.2025-0034

Do, T. B. T. (2025b). Teachers' experiences and perceptions of using ChatGPT in writing instruction. *VNU Journal of Foreign Studies*, *41*(1S), 180-197. https://doi.org/10.63023/2525-2445/jfs.ulis.5473

Dündar, O. İ. et Aydoğdu, C. (2024). Création d'un agent conversationnel pour une classe de FLE : les apports et les limites pour l'expression écrite. *Anadolu Journal of Educational Sciences International*, *14*(2), 501-523. https://doi.org/10.18039/ajesi.1411872

Elmouhtadi, A., Souhad, S. et Soughati, N. (2025). L'intelligence artificielle générative (IAG) dans l'apprentissage du FLE chez l'étudiante et l'étudiant marocains : étude de cas de la Faculté des sciences de Rabat. *Revue internationale des technologies en pédagogie universitaire / International Journal of Technologies in Higher Education*, *22*(1), Article 16. https://doi.org/10.18162/ritpu-2025-v22n1-16

El Wahabi, S. et Belhaj, L. (2024). Preparing learners for the higher education curriculum: Perfecting argumentative writing skills in FFL—An innovative teaching approach with ChatGPT and interactive in-class assessment. *International Journal of Sciences: Basic and Applied Research*, *73*(1), 31-42.

Ferris, D. R. (2012). Written corrective feedback in second language acquisition and writing studies. *Language Teaching, 45*(4), 446-459. https://doi.org/10.1017/S0261444812000250

Godwin-Jones, R. (2024). Distributed agency in language learning and teaching through generative AI. *Language Learning & Technology, 28*(2), 5-30. https://doi.org/10.64152/10125/73570

Hayes, J. R. (2012). *Modeling and remodeling writing*. *Written Communication, 29*(3), 369-388. https://doi.org/10.1177/0741088312451260

Heddouche, O. (2024). L'agent conversationnel ChatGPT : un tuteur virtuel au service du développement des compétences rédactionnelles en classe de FLE. *Crossing Boundaries in Culture and Communication*, *15*(3), 107-116.

Holmes, T. (2024). *ChatGPT pour la rétroaction corrective écrite interactive en apprentissage du français langue seconde* [Mémoire de maîtrise, Université d'Ottawa]. uO Research. https://doi.org/10.20381/ruor-30407

Khelafi, S. et Bourkaib Saci, N. (2024). L'utilisation de ChatGPT-3 pour le développement des stratégies argumentatives dans la rédaction d'articles scientifiques en FLE chez les chercheurs algériens. *Milev Journal of Research & Studies*, *10*(2), 232-253. https://asjp.cerist.dz/en/article/263703

Levac, D., Colquhoun, H. et O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, *5*, 69. https://doi.org/10.1186/1748-5908-5-69

Li, Y. (2025). L'intelligence artificielle générative dans l'apprentissage du français en Chine : une étude sur l'utilisation de ChatGPT pour le développement des compétences orthographiques. *Didactique du FLES*, *4*(1). https://doi.org/10.57086/dfles.1608

Lorrillard, O. (2025). Intelligence artificielle générative et apprentissage du français : une étude de cas au Japon. *Didactique du FLES*, *4*(1), 49-68. https://doi.org/10.57086/dfles.1591

Marchal, B. et Phoungsub, M. (2025). L'intelligence artificielle générative dans l'enseignement du FLE en Asie du Sud-Est : défis éthiques vers une rédaction créative computationnelle. *Bulletin de l'ATPF*, (149), 53-66.

Mehrabi, M. et Haghgou, Z. (2025). L'impact de ChatGPT sur la production écrite des étudiants iraniens : une analyse d'erreurs. *Didactique du FLES*, *4*(1). https://doi.org/10.57086/dfles.1583

Molina Zapata, V. et Montoya Estrada, J. D. (2025). I.A. améliorer l'expérience FLE : L'impact de l'intégration de l'intelligence artificielle dans l'apprentissage du FLE chez un groupe d'adolescentes à l'Alliance française de Medellín. *Didactique du FLES*, *4*(1). https://doi.org/10.57086/dfles.1613

Nieto Escobar, M. et Nadim, L. (2024). Les opportunités d'enseignement/apprentissage du français langue étrangère avec ChatGPT à l'ère de l'intelligence artificielle et de l'apprentissage mobile. *Synergies Espagne*, *17*, 237-258.

Rinck, F. (2025). Des textes de l'IAG pour former à l'enseignement de la production écrite. *Revue internationale des technologies en pédagogie universitaire / International Journal of Technologies in Higher Education*, *22*(1), Article 10. https://doi.org/10.18162/ritpu-2025-v22n1-10

Schlemminger, G. (2025a). Apprentissage du français en migration : l'IAG pour plus d'autonomie. *Didactique du FLES*, *4*(1). https://doi.org/10.57086/dfles.1624

Schlemminger, G. (2025b). L'intelligence artificielle générative au service du FLES, à l'exemple de ChatGPT. *Didactique du FLES*, *4*(1). https://doi.org/10.57086/dfles.1573

Schlemminger, G. et Gettliffe, N. (2025). Usages raisonnés de l'intelligence artificielle générative (IAG) pour l'enseignement du français langue étrangère et seconde. *Didactique du FLES*, *4*(1), 3-9. https://www.ouvroir.fr/dfles/index.php?id=1571

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. Annals of Internal Medicine, *169*(7), 467-473. https://doi.org/10.7326/M18-0850

Warschauer, M., Tseng, W., Yim, S., Webster, T., Jacob, S., Du, Q. et Tate, T. (2023). The affordances and contradictions of AI-generated text for second language writers. *Journal of Second Language Writing, 62*, 101071. https://doi.org/10.1016/j.jslw.2023.101071

Annexes

#	Citation (APA7)	Année	Pays / Contexte	Population / Niveau CECR	Type d'apprenants	Outil IAg	Tâches d'écriture	Conception de l'étude	Durée / Contexte pédagogique	Indicateurs / Mesures	Principaux résultats	Enjeux éthiques / Limites	Notes
3	Holmes, T. (2024). *ChatGPT pour la rétroaction corrective écrite interactive en apprentissage du français langue seconde* [Mémoire de maîtrise, Université d'Ottawa]. uO Research. https://doi.org/10.20381/ruor-30407	2024	Canada ; université (cours de FLS)	n = 22, étudiants FLS, niveau "intermédiaire faible".	Étudiants/apprenants en FLS (contexte universitaire).	ChatGPT.	Soumission d'un texte en français suivie d'échanges apprenant-ChatGPT pour rétroaction corrective écrite interactive (RCÉ).	Étude empirique expérimentale avec analyse taxonomique des fils d'interactions ; questionnaires pré-/post-tâche.	Intervention ponctuelle au sein d'un cours universitaire (durée exacte non précisée).	Typologie des requêtes étudiantes ; typologie/ qualité des réponses de ChatGPT ; perceptions (pré/post). Tableaux 4-7 ; Figure 22.	Réponses de ChatGPT majoritairement correctes et appropriées ; perceptions positives des participants ; qualité variable et pouvant baisser si la requête est imprécise (exemples d'erreurs et de réponses inappropriées documentés).	Limites illustrées par des réponses incorrectes (p. ex. erreurs de conjugaison, modifications non réalisées) et des réponses inappropriées ; dépendance à la qualité du prompt. (Fig. 23-28 montrent ces cas.)	Contenu riche : Tableaux 4-7 (données quantitatives, typologies des requêtes/réponses), Figure 22 (alignement requête-réponse).
4	Elmouhtadi, A., Souhad, S. et Soughati, N. (2025). L'intelligence artificielle générative (IAG) dans l'apprentissage du FLE chez l'étudiante et l'étudiant marocains : étude de cas de la Faculté des sciences de Rabat. *Revue internationale des technologies en pédagogie universitaire / International Journal of Technologies in Higher Education*, *22*(1), Article 16. https://doi.org/10.18162/ritpu-2025-v22n1-16	2025	Maroc ; université (Faculté des sciences de Rabat, FSR).	Étudiants de première année, niveau débutant (A1-A2) (CECR).	Étudiant·e·s universitaires des filières scientifiques MIP et InfApp.	ChatGPT, l'étude parle plus largement d'IAG.	Quatre séances de production écrite en présentiel ; le recours à des tuteurs IAG est autorisé et la consigne précise : « Vous pouvez utiliser ChatGPT dans votre production écrite. » Les tâches visent l'organisation des idées, la correction grammaticale/ syntaxique et le lexique.	Méthodologie mixte : approche quantitative (enquête par questionnaire) + approche qualitative (observation participante en cours de FLE) ; ancrage recherche-action/ingénierie pédagogique.	Dispositif hybride (présence + autonomie) : chaque semaine 1 h 30 de cours en présentiel, poursuite du travail sur Moodle ; activités en ligne synchrones (Google Meet) et asynchrones.	Enquête quantitative auprès de la population cible (500 étudiants) avec 316 réponses valides (44 %) ; items sur usages de l'IAg, perceptions, autonomie, esprit critique/éthique ; observation des pratiques d'écriture (intégration des suggestions IAG).	L'IAG est largement présente dans les pratiques ; elle soutient la coconstruction des compétences linguistiques et le développement de l'autonomie ainsi que de l'esprit critique et éthique chez une grande partie des apprenants qui l'utilisent de manière réfléchie.	L'article met l'accent sur la littératie éthique et critique vis-à-vis de l'IAG (usage réfléchi) ; n'indique pas d'expérimentation contrôlée versus témoin, ni de mesure de gains d'apprentissage standardisée – limite pour l'attribution causale.	Le cours cible A1-A2 (première année) en FLE scientifique (MIP/ InfApp). Questionnaire en 46 questions (AR, majoritairement fermées) ; autorisation explicite d'usage de ChatGPT pendant les 4 séances de production écrite. Dispositif pensé pour favoriser l'autonomie d'écriture dans un cadre hybride.
5	Mehrabi, M. et Haghgou, Z. (2025). L'impact de ChatGPT sur la production écrite des étudiants iraniens : une analyse d'erreurs. *Didactique du FLES*, *4*(1). https://doi.org/10.57086/dfles.1583	2025	Iran ; Université de Téhéran, master à distance en didactique du FLE (contexte décrit en 3.1).	17 étudiants inscrits au cours ciblé ; 13 ont été inclus pour l'étude (3.2) ; 10 ont mené l'intervention jusqu'au bout (poursuite avec 10 restants). Niveau des admis généralement B1-B2 CECR ; l'étude cible 13 étudiants de niveau B1 (3.1).	Étudiants de master (FLE) en formation à distance.	ChatGPT (outil étudié tout au long de l'intervention).	Pré-test et post-test composés de 3 sujets chacun, « adaptés à leur niveau B1 », réalisés avant et après l'intervention de ChatGPT. Les thèmes couvrent : communication professionnelle (lettres), réflexion/opinion (ex. réseaux sociaux, éducation environnementale), récit d'expérience (3.3 ; description détaillée des 3 catégories et exemples).	Approche mixte : analyse quantitative des scores pré/post (tests de normalité Kolmogorov-Smirnov, t apparié), graphique des moyennes ; analyse qualitative des productions selon 4 opérations (omission, addition, substitution, déplacement) et classification des erreurs ; entretiens thématiques (sections 4-6).	Étude menée pendant un semestre (3.1). Séquence d'évaluation pré-/post-autour de l'intervention avec ChatGPT ; formation en distanciel (programme de master à distance).	Scores aux pré/post-tests (Tableau des notes ; tests K-S et t apparié), moyennes comparées (Graphique 1) ; typologie d'erreurs (Tableau de classification) et exemples annotés ; entretiens (perceptions positives/ négatives ; sections 6.1-6.4).	Amélioration significative des scores entre pré-test et post-test (normalité vérifiée ; t apparié, $p < .05$ pour les trois sujets) ; ChatGPT perçu comme partenaire utile pour la correction des erreurs et la révision ; mise en évidence des catégories d'erreurs dominantes avant intervention et d'une réduction post-intervention (sections 4-6 ; Graphique 1 ; Tableaux K-S/t).	Taille d'échantillon réduite (10 complétants) ; contexte à distance avec accès restreint à ChatGPT en Iran (contraintes d'inscription/ connexion) → limite d'équité d'accès ; pas de groupe témoin simultané, donc attribution causale prudente (sections contexte/discussion).	Détaille soigneusement : consignes des 3 catégories de sujets, Tableau des notes pré/post, tests K-S et t, Graphique des moyennes, tableau de classification des erreurs ; sections 6.1-6.4 synthétisent impact, efficacité comme partenaire de correction, perceptions +/-.
6	Li, Y. (2025). L'intelligence artificielle générative dans l'apprentissage du français en Chine : une étude sur l'utilisation de ChatGPT pour le développement des compétences orthographiques. *Didactique du FLES*, *4*(1). https://doi.org/10.57086/dfles.1608	2025	Chine ; université (licence FLE).	Deux groupes de 14 étudiants ; groupe expérimental vs témoin ; niveau ≈ B1 compréhension écrite et A2/B1 production écrite (p. 7-8).	Étudiants de 2e année de licence FLE (p. 7).	ChatGPT (p. 2 : plan ; p. 8 : mise en œuvre).	Dictée dialoguée + entretiens métalinguistiques ; préparer des requêtes précises ; analyse et justification des corrections (p. 8-10).	Quasi-expérimentale (groupe expérimental/ témoin) + approche qualitative (entretiens métalinguistiques) ; pré-/post- par dictées évaluatives (p. 7-8).	6 semaines, 1 dictée dialoguée/sem. ; pré/post par dictée traditionnelle (p. 8).	Comptage de l'emploi du métalangage et des manipulations syntaxiques ; suivi d'erreurs et tâches personnalisées (p. 14-15).	IAG soutient la verbalisation et l'autorévision ; vers une utilisation autonome avec suivi pédagogique adapté (p. 7).	Met l'accent sur le développement de protocoles d'utilisation et les suivis pédagogiques; pas de vérification contrôlée de longue durée.	Section « Consignes données à ChatGPT » (exemples de prompts) (p. 2, 9). Ex. : « Peux-tu corriger mon texte… ? » ; « Peux-tu expliquer pourquoi… ? » (p. 9).
10	Schlemminger, G. (2025a). Apprentissage du français en migration : l'IA générative pour plus d'autonomie. Didactique du FLES, 4(1). https://www.ouvroir.fr/dfles/index.php?id=1624 ; DOI: 10.57086/dfles.1624.	2025	Contexte migratoire (centre associatif) en FLE.	≈ 10 adultes A1, demandeur·euses d'asile.	Adultes allophones débutants (A1) en situation de migration.	ChatGPT.	Génération et révision d'énoncés courts (vocabulaire/vêtements), rédaction de texte personnel, dialogue en magasin ; formulation de prompts.	Recherche-action en pédagogie active, itérative, combinant observation, ajustements pédagogiques et évaluation des acquis.	Non précisé (nombre de séances/semaines non indiqué) ; atelier en présentiel avec appuis outils numériques (ex. FLElex).	Suivi qualitatif des productions et de la littératie IAG (prompting, exploitation des réponses), appuyé par outils d'analyse lexicale (FLElex).	L'IAG soutient l'autonomie, la créativité et la conscience linguistique tout en facilitant l'appropriation lexicale/syntaxique.	Nécessité d'une littératie numérique en IA et d'un guidage pour éviter dépendance/erreurs ; accès restreint (quotas).	Présence d'exemples de prompts (FR/EN), d'outputs (textes/ dialogues) et d'un ancrage Freinet (texte libre).
11	Molina Zapata, V. et Montoya Estrada, J. D. (2025). I.A. améliorer l'expérience FLE : L'impact de l'intégration de l'intelligence artificielle dans l'apprentissage du FLE chez un groupe d'adolescentes à l'Alliance française de Medellín. *Didactique du FLES*, *4*(1). https://doi.org/10.57086/dfles.1613.	2025	Colombie ; Alliance française de Medellín (AFM).	20 adolescentes bénéficiaires de bourse; niveaux organisés en sous-niveaux de 40h selon le CECRL (niveau précis non indiqué).	Adolescentes en cours FLE à l'AFM (formation présentielle). Contexte décrit.	Application d'IA pour entraînement linguistique ; génération de chansons avec IA SUNO ; travail de prompts.	Production écrite évaluée (pré/post); rédaction de prompts; activités créatives (paroles de chansons) adossées à objectifs d'apprentissage.	Recherche-action avec pré-test/post-test (modèle DELF) + enquête de perceptions; groupe expérimental vs groupe contrôle.	Durée exacte non précisée; trois interventions décrites (initiation à l'IA & éthique; jeu en ligne de caractéristiques d'un bon prompt; activités créatives).	Scores CE, PE, PO, NF (avant/après) ; graphiques comparatifs; entretiens et enquêtes.	Écrit : le groupe contrôle obtient de meilleures performances en production écrite; le groupe expérimental progresse davantage en compréhension écrite et production orale.	Risques : dépendance, erreurs/approximations; nécessité d'un encadrement et d'une utilisation équilibrée.	Mots-clés : « intelligence artificielle, FLE, TICE, motivation, autonomie » ; exemple sur stéréotypes générés par IA (besoin d'accompagnement enseignant).

#	Citation (APA7)	Année	Pays / Contexte	Population / Niveau CECR	Type d'apprenants	Outil IAg	Tâches d'écriture	Conception de l'étude	Durée / Contexte pédagogique	Indicateurs / Mesures	Principaux résultats	Enjeux éthiques / Limites	Notes
12	Marchal, B. et Phoungsub, M. (2025). L'intelligence artificielle générative dans l'enseignement du FLE en Asie du Sud-Est : défis éthiques vers une rédaction créative computationnelle. *Bulletin de l'ATPF*, (149), 53-66.	2025	Thaïlande (universités Thammasat & Chiang Maï) ; étude centrée FLE en Asie du Sud-Est (p. 27, 36).	N = 124 étudiants (69 Thammasat, 55 Chiang Maï) ; niveaux A1 3,2 % ; A2 16,9 % ; B1 61,3 % ; B2 17,7 % ; C1 0,8 %.	Étudiants licence (3e-5e année) FLE en Thaïlande.	Outils d'IA incluant les robots conversationnels (ex. ChatGPT), traducteurs, correcteurs.	Usages déclarés de l'IAG pour la production écrite (correction, enrichissement lexical, révision/autocorrection).	Enquête par questionnaire (Google Forms), majoritairement quantitative (questions fermées), validée par 2 enseignants-chercheurs + prétest.	Collecte durant 1 mois – septembre 2024.	Fréquence d'usage, types d'outils, difficultés en écrit, attentes/utilité perçue pour l'écrit.	Usage fréquent de l'IA pour l'écrit ; 84,1 % utilisent des robots conversationnels ; attentes élevées pour correction (92,9 %), lexique (77 %) et révision (68,1 %).	Risques : fraude, dépendance, incohérences hallucinatoires ; besoin de cadres éthiques et formation.	Résumé FR explicite le périmètre : « explore les usages… dans les productions écrites » (p. 27). Figures synthétiques (p. 38). Contexte éthique/ politique détaillé (p. 32-33, 49).
13	El Wahabi, S. et Belhaj, L. (2024). Preparing learners for the higher education curriculum: Perfecting argumentative writing skills in FFL—An innovative teaching approach with ChatGPT and interactive in-class assessment. *International Journal of Sciences: Basic and Applied Research*, *73*(1), 31-42.	2024	Maroc ; lycée (voie sciences expérimentales), Homman El Fatwaki.	n = 82 élèves de première année du baccalauréat. Niveau CECR : non précisé.	Adolescents scolarisés au lycée (voie sciences expérimentales).	ChatGPT.	Essais argumentatifs réguliers (trouver des arguments, connecteurs logiques) + feedback en classe.	Approche mixte (qualitative + quantitative), intervention pédagogique avec évaluations en classe et examen régional.	4 mois : un essai chaque week-end, + textes supplémentaires pendant vacances ; accélération en mai (3/semaine).	Qualité/cohérence des textes ; usage des arguments et connecteurs ; introductions/conclusions ; notes d'évaluations en classe et examen régional.	Amélioration notable d'une part importante des élèves ; 41,46 % « succès significatif ».	Échantillon unique, durée 4 mois, généralisabilité limitée ; besoins de formation/encadrement éthique.	Leçons d'argumentation avant usage d'IA ; évaluation interactive (lecture à voix haute, feedback pair/ enseignant).
14	Khelafi, S. et Bourkaib Saci, N. (2024). L'utilisation de ChatGPT-3 pour le développement des stratégies argumentatives dans la rédaction d'articles scientifiques en FLE chez les *chercheurs algériens. Milev Journal of Research & Studies*, *10*(2), 232-253. https://asjp.cerist.dz/en/article/263703	2024	Algérie ; rédaction d'articles scientifiques en FLE par des chercheurs.	Chercheurs algériens (effectif non précisé) sollicités pour intégrer ChatGPT-3 dans leur processus de rédaction. Niveau CECR : non applicable/non précisé (public chercheur, non scolarisé CECR).	Adultes chercheurs (non-étudiants). Contexte de rédaction scientifique (p. 1).	ChatGPT-3.	Production/révision de passages d'articles scientifiques avec ciblage de stratégies argumentatives (concession, réfutation, explication) et prompting.	Étude qualitative d'interactions chercheur-ChatGPT-3 + comparaison versions « originale » vs « retravaillée ».	Non précisé (pas de calendrier de séances). Travail hors classe sur manuscrits scientifiques (p. 1, 6) ; focalisé sur la phase de rédaction et le prompt engineering (p. 6).	Grille d'évaluation incluant : « Clarté de la réponse », « Pertinence de la réponse », « Qualité de l'argumentation », « Orthographe et grammaire »	Amélioration notable de la qualité de l'argumentation dans les versions retravaillées ; capacités et limites de l'IA mises en évidence.	Pas exempt d'erreurs ; risque de redondances ; dépend de la clarté des demandes.	Démonstrations détaillées par cas : Recherche 1/2/3 (p. 12, 14, 16) ; exemples de connecteurs logiques et d'appels à l'autorité / réfutation / explication (p. 13-16, 19-21). Prompt engineering défini et mobilisé (p. 6).
15	Bechiri, C. (2024). Intégration de l'intelligence artificielle dans la classe de FLE : approches et pratiques pour l'amélioration de l'écrit à l'université de Skikda. *Ziglôbitha – Revue des Arts, Linguistique, Littérature & Civilisations*, *3*(10), 139-148.	2024	Algérie ; Université de Skikda (contexte universitaire).	Étudiants de master (niveau CECR non précisé).	Universitaires (master) en FLE.	ChatGPT et Perplexity.	Dissertation sur un thème didactique (écrit vs oral).	Trois cohortes : 2 avec IA (ChatGPT/Perplexity), 1 sans IA (témoin) ; 3 séances intensives (3h chacune).	3 séances intensives × 3h dans un cours de didactique de l'écrit.	Qualité rédactionnelle observée : structure et erreurs linguistiques.	Amélioration significative pour groupes IA ; groupe sans IA en difficulté.	Risque d'homogénéisation et dépendance ; besoin d'encadrement enseignant.	Détails de publication : Ziglôbitha (ISSN-L 2708-390X) (p. 1) ; mention des outils (Google Bard, Perplexity, Claude, ChatGPT) (p. 3) ; ancrage pédagogique et rôle de l'enseignant (p. 6-7).
16	Dündar, O. İ. et Aydoğdu, C. (2024). Création d'un agent conversationnel pour une classe de FLE : les apports et les limites pour l'expression écrite. *Anadolu Journal of Educational Sciences International*, *14*(2), 501-523. https://doi.org/10.18039/ajesi.1411872	2024	Turquie ; Université Anadolu, Département de FLE (licence, Pédagogie).	n = 12 apprenants ; niveau B1+.	Étudiant·e·s de 1re année de licence inscrits au cours « Expression écrite II ».	Agents conversationnels (AC) conçus par les auteur·e·s (8 bots, logique d'arbre décisionnel).	Séries de tâches d'expression écrite alignées B1+, avec autoévaluation à chaque tâche.	Méthode mixte : recherche-développement + « modèle chronologique » ; questionnaire initial/final et analyses de journaux/ tâches.	Semestre de printemps 2021-2022 ; cours présentiel, possibilité d'usage à distance en cas d'absence.	volution des stratégies d'écriture (avant/pendant/ après) via questionnaire initial/final ; grille DELF adaptée en questions fermées ; analyse des autoévaluations par tâche.	Effet significatif sur le développement de l'expression écrite ; hausse des stratégies liées au plan d'écriture ; appropriation des critères DELF.	Échantillon restreint (n=12) ; dispositif limité à B1+ ; points à réviser pour des usages futurs ; accès/usage à distance ponctuel.	Les 8 AC intègrent des consignes alignées CECRL ; autoévaluations répétées par tâche ; items détaillés (11 questions) couvrant adéquation, enchaînement, lexique, orthographe/grammaire (p. 508-509).
17	Do, T. B. T. (2025a). Feedback quality of ChatGPT in revising French texts. *HNUE Journal of Science: Educational Sciences*, *70*(2), 33-44. https://doi.org/10.18173/2354-1075.2025-0034	2025	Viêt Nam ; ULIS-VNU Hà Nội (université).	Corpus de 20 textes d'apprenants A2+.	Étudiants FLE (A2+) – contexte universitaire.	ChatGPT.	Révision de textes : détection d'erreurs, analyse d'erreurs, proposition de correction.	Méthodes mixtes sur corpus de 20 textes : codage des erreurs et évaluation de la qualité des retours.	Non précisé (étude de corpus hors intervention chronométrée).	Taux de détection erronée, analyse erronée, proposition erronée ; typologie d'erreurs & cas de surcorrection.	Qualité hétérogène : parts de surcorrection > 50 % ; 29 % des mauvaises analyses dues à un manque de connaissance métalinguistique.	Limites d'interprétation contextuelle et risque de dépendance au prompt ; besoin de préciser l'intention/lecteur pour orienter l'IA.	Tableaux de codage & exemples annotés (p. 5, 9-10). Extrait (p. 5) : « catégories Correct/ Incorrect pour détection/ analyse/correction ».
19	Heddouche, O. (2024). L'agent conversationnel ChatGPT : un tuteur virtuel au service du développement des compétences rédactionnelles en classe de FLE. *Crossing Boundaries in Culture and Communication*, *15*(3), 107-116.	2024	Algérie ; Université de Biskra ; module de pratiques rédactionnelles (Master).	Étudiants de master (effectif non précisé).	Universitaires (Master FLE).	ChatGPT (agent conversationnel/tuteur virtuel).	Résumé d'un texte → identifier/analyser les erreurs via ChatGPT → révision du résumé.	Étude qualitative avec comparaison texte initial vs texte révisé ; protocole en trois phases.	Non précisé (pas de nombre de séances/ semaines). Contexte : cours de Master – pratiques rédactionnelles.	Comparaison qualitative de la qualité des résumés (avant/après), autonomie dans l'analyse des erreurs, retours des étudiants.	Impact positif sur progression et motivation ; amélioration de la qualité des résumés ; autonomie accrue.	Peut produire des réponses incorrectes ; besoin de littératie numérique/éthique et d'un encadrement.	Mots-clés : « compétences rédactionnelles … motivation … » ; Perspectives pédagogiques : autoévaluation, personnalisation, collaboration humain-IA.

#	Citation (APA7)	Année	Pays / Contexte	Population / Niveau CECR	Type d'apprenants	Outil IAg	Tâches d'écriture	Conception de l'étude	Durée / Contexte pédagogique	Indicateurs / Mesures	Principaux résultats	Enjeux éthiques / Limites
7	Lorrillard, O. (2025). Intelligence artificielle générative et apprentissage du français : une étude de cas au Japon. *Didactique du FLES*, *4*(1), 49-68. https://doi.org/10.57086/dfles.1591	2025	Japon ; étude de cas sur l'usage de l'IAG en contexte FLE au Japon.	Débutants.	Jeunes apprenants japonais débutants.	ChatGPT.	Échanges de correction/ révision d'énoncés ; analyse des erreurs et des explications de l'IAG.	Retour d'expérience / étude de cas avec protocole d'évaluation des réponses de l'IAG.	Non précisé.	Catégories d'erreurs (orthographe, traduction, « sans contexte », « en contexte », ponctuation) et qualité des explications.	L'analyse met en évidence des corrections inégales, des explications parfois hors-sujet/erronées, des corrections excessives et des tensions avec la simplification.	Étude sans groupe témoin ni pré/post explicités ; focalisée sur l'évaluation des sorties IAG en contexte débutant → limites d'attribution causale (aucune procédure expérimentale comparative décrite ; p. 2, structure méthodo/résultats).
8	Schlemminger, G. (2025b). L'intelligence artificielle générative au service du FLES, à l'exemple de ChatGPT. *Didactique du FLES*, 4(1). https://doi.org/10.57086/dfles.1573	2025	Contexte FLES ; revue scientifique francophone. Article de point de vue.	Non applicable (article de point de vue, pas d'échantillon).	Non applicable.	ChatGPT.	Non applicable (pas d'étude empirique sur des tâches).	Article de synthèse / point de vue structuré par rubriques : fonctionnement, questions épistémologiques/éthiques, limites, conseils d'usage.	Non précisé (pas d'intervention pédagogique).	Non applicable (pas de mesures empiriques).	L'article interroge les apports, limites et implications éthiques de l'IAG en FLES et propose des pistes d'usage raisonné.	Souligne des « limites… sur le plan épistémologique qu'éthique » et les risques de substitution de l'interaction humaine.
9	Schlemminger, G. et Gettliffe, N. (2025). Usages raisonnés de l'intelligence artificielle générative (IAG) pour l'enseignement du français langue étrangère et seconde. *Didactique du FLES*, *4*(1). https://www.ouvroir.fr/dfles/index.php?id=1571	2025	Contexte général FLE/FLES ; article éditorial du numéro thématique *L'intelligence artificielle générative pour l'enseignement du FLE*. Minh chứng ; titre du numéro : « L'intelligence artificielle générative pour l'enseignement du FLE » ; section indiquée : « Éditorial ».	Non applicable : l'article ne présente pas de population d'étude propre ni de niveau CECR spécifique. Il introduit les contributions du numéro.	Non applicable : aucun échantillon d'apprenants n'est constitué dans cet éditorial. L'article mentionne seulement les publics étudiés dans d'autres contributions du numéro.	IAG / ChatGPT, envisagés de manière générale.	Non applicable comme tâche expérimentale propre.	Éditorial / introduction d'un numéro spécial ; pas d'étude empirique propre.	Aucun dispositif pédagogique propre n'est décrit.	Non applicable : aucun indicateur ni mesure propre à cet article.	L'article propose une synthèse programmatique : le numéro montre la diversité des usages de l'IAG en didactique du FLES et appelle à des usages humanistes, critiques et émancipateurs.	L'article souligne plusieurs enjeux : absence de consentement, manque de traçabilité/transparence, biais culturel, contenus discriminatoires, consommation énergétique, durabilité et régulation.
18	Do, T. B. T. (2025b). Teachers' experiences and perceptions of using ChatGPT in writing instruction. *VNU Journal of Foreign Studies, 41*(1S), 180-197. https://doi.org/10.63023/2525-2445/jfs.ulis.5473	2025	Viêt Nam ; enseignants de FLE au lycée et à l'université.	53 enseignant·e·s de FLE (utilisateurs de ChatGPT en cours de rédaction). CECR : non applicable (public enseignants).	Enseignant·e·s de FLE (échantillon d'enseignants ; pas d'apprenants).	ChatGPT (objet d'étude).	Focalisation sur la rédaction de dissertation : usages rapportés pour générer des plans, exemples, matériels de lecture ; intégrer la consigne dans le prompt ; révision basée sur le feedback.	Enquête quantitative (questionnaire 42 items), statistiques descriptives (SPSS 26).	Période de collecte : non précisée dans le texte ; article accepté en mai 2025. (Indications éditoriales seulement.)	Fréquences moyennes (Likert) par volets Planification / Mise en œuvre / Évaluation (Tableaux 1-3). Exemples (moyennes) : consigne dans le prompt (M=3.30), corriger erreurs non détectées (M=3.28).	Usages fréquents avant cours (plan, exemples, lectures) ; en classe, limiter la dépendance et autoriser le plan ; en évaluation, combiner prompt + correction humaine ; bénéfices (gain de temps, feedback instantané) ; risques (triche académique, inexactitude).	Préoccupations majeures : malhonnêteté académique et inexactitude du feedback IA ; besoin de formation (prompting, grammaire/ vocabulaire assistés par IA).
20	Constantin, F. (2023). ChatGPT – Learning accelerator or demolisher of foreign language teaching and learning? An empirical study on *Business French. The Annals of the University of Oradea, Economic Sciences*, *32*(2), 225-238. https://doi.org/10.47535/1991auoes32(2)022	2023	Roumanie ; Université d'Oradea ; Business French (FOS).	Non précisé.	Étudiants Business French (niveau non indiqué).	ChatGPT.	Lettre de vente en deux temps : (I) élaboration humaine (LM → FR) ; (II) génération par ChatGPT de versions FR multiples.	Scénario pédagogique démonstratif, en 2 phases pour contraster HI vs IA, sans pré/ post-tests standardisés.	2 heures.	Observations qualitatives : rapidité, nombre de versions, adéquation au style professionnel (pas de métriques chiffrées d'apprentissage).	Effet de choc didactique, prise de recul critique ; nécessité d'expertise pour évaluer la correction.	Problèmes d'évaluation et risque de dépendance/ contournement ; absence de données apprenants et de mesures d'efficacité.
21	Nieto Escobar, M. et Nadim, L. (2024). Les opportunités d'enseignement/apprentissage du français langue étrangère avec ChatGPT à l'ère de l'intelligence artificielle et de l'apprentissage mobile. *Synergies Espagne*, *17*, 237-258.	2024	Espagne; cadre FLE universitaire ; approche m-learning + ChatGPT.	Non précisé / non applicable (article de conception pédagogique, sans échantillon).	Non précisé (ciblage large FLE ; exemples au niveau B1 pour activités).	ChatGPT (robot conversationnel) comme outil didactique central.	Production écrite (catalogue d'activités) : générer texte + QCM/QA, résumer, complétion de texte, correction/amélioration, brouillon, idées/arguments. Exemples d'invites (B1 CECR)	Article de cadrage / conception pédagogique présentant un catalogue d'activités + analyse critique.	Non précisé (pas d'intervention chronométrée).	Non applicable (pas de mesures empiriques).	Positionnement : ne pas diaboliser ; intégrer l'IA avec compétence numérique et guidage ; IA = alliée si orientée vers compétences.	Limites : imprécisions, dépendance, corriger sans localiser l'erreur ; risques droits d'auteur, fiabilité/ sources, coût, empreinte énergétique.
22	Rinck, F. (2025). Des textes de l'IAG pour former à l'enseignement de la production écrite. *Revue internationale des technologies en pédagogie universitaire / International Journal of Technologies in Higher Education*, *22*(1), Article 10. https://doi.org/10.18162/ritpu-2025-v22n1-10	2025	France ; formation initiale d'enseignant·e·s du primaire (didactique de l'écriture).	Stagiaires enseignant·e·s du primaire (pas d'apprenants scolaires analysés en pré/post) – CECR : non applicable.	Enseignant·e·s en formation (public adulte, formation initiale).	GPT-4 / ChatGPT (textes générés).	Récit narratif à partir d'un lanceur (« *Ils fouillent le grenier… 38 000 euros* ») ; rédaction, comparaison de corpus (élèves, stagiaires, IA), réécriture.	Expérience de formation : travail en groupes (4-5) pour formuler/hiérarchiser des critères de réussite puis critères de réalisation ; mises en commun guidées.	Non précisé (séquence de formation universitaire ; projet Pégase – U. Grenoble Alpes).	Indicia qualitatifs : élaboration de critères (univers de référence / énigme / états mentaux), réécriture d'un texte d'élève, échanges réflexifs.	Prises de conscience sur la textualité et le processus d'écriture ; les textes IA comme « copilote utile » pour outiller l'étayage.	Pas d'évaluation chiffrée d'apprenants ; certains jugent les textes IA stéréotypés/peu savoureux ; nécessité de ne pas viser la conformité au modèle IA.